## DOCTOR OF PHILOSOPHY

# Identification and characterization of the gene responsible for the matted mouse phenotype

Goh, Christabelle

*Award date:*
2013

DOCTOR OF PHILOSOPHY

# Identification and characterization of the gene responsible for the matted mouse phenotype

Christabelle Goh

2013

University of Dundee

# Identification and characterization of the gene responsible for the *matted* mouse phenotype

**Christabelle Goh Si Mei**

**Thesis presented for the degree of**

**Doctor of Philosophy**

**University of Dundee**

**October 2013**

# TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# ACKNOWLEDGEMENTS

It takes a village to raise a child, in much the same way, this thesis would never have been written without the guidance and support from the entire McLean lab. WHIM FTW!

I am grateful for Professor Irwin McLean offering me the opportunity to pursue my Phd, and for having the confidence and conviction that I could pull it off. I am better off for experiencing your mentorship and friendship.

Dr. Aileen Sandilands is possibly the world's best supervisor, and I am so lucky and thankful to have her as mine.

Much love to Frances, Emma, Liz and Linda for all their generosity and advice; Sara for all her help with statistics; everyone else in the lab for their company.

Many thanks to Chris Cole, Andy Prescott and John James and all their help with the technical bits of this project.

Of course I would never have been able to go through this ordeal without all the support and love from my soulmate. Thank you for being the anchor while I weathered this storm.

Much love to my parents for all the prayers, concern and recipes.

# DECLARATION

I declare that I am the sole author of this thesis and that all references cited have been consulted by me personally. The work, of which this thesis is a record, has been done by myself, unless otherwise acknowledged. This work has not been previously submitted for a higher degree.

Signed ……..

**Christabelle Goh Si Mei**

Date ……

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AD | Atopic dermatitis |
| ARC | Arthrogryposis, renal dysfunction and cholestasis |
| BKC | Blepharokeratoconjunctivitis |
| CTSC | Cathepsin C |
| CDSN | Corneodesmosin |
| CE | Cornified envelope |
| CDS | Chanarin-Dorfman syndrome |
| cl | Companion layer |
| co | Cortex |
| cu | Cuticle layer |
| DM | Double mutant |
| DMEM | Dulbecco's Modified Eagle Medium |
| DP | Dermal papilla |
| DSC | Desmocollin |
| DSG | Desmoglein |
| DSP | Desmoplakin |
| EBS | Epidermolysis bullosa simplex |
| EDC | Epidermal differentiation complex |
| EEA | Early endosome antigen |
| EHK | Epidermolytic hyperkeratosis |
| EPPK | Epidermal palmoplantar keratoderma |
| ESCs | Epidermal stem cells |
| FCS | Fetal calf serum |
| FLG | Filaggrin |

| | |
|---|---|
| GAPDH | Glyceraldehyde 3-phosphate dehydrogenase |
| GFP | Green fluorescent protein |
| He | Henle layer |
| HI | Harlequin ichthyosis |
| HPK | Human primary keratinocytes |
| HS | Hair shaft |
| Hu | Huxley layer |
| icu | Inner cuticle |
| InDel | Insertion/Deletion |
| INV | Involucrin |
| IRS | Inner root sheath |
| IV | *Ichthyosis vulgaris* |
| IVL | Involucrin |
| KIF | Keratin intermediate filaments |
| KLK | Kallikrein |
| LAMP | Lysosomal-associated membrane protein |
| LCE | Late cornified envelope |
| LCR | Locus control region |
| LEKTI | Lymphyo-epithelial Kazal-type-related inhibitor |
| LG | Lamellar granules |
| LI | Lamellar ichthyosis |
| LOR | Loricrin |
| ma | Matrix |
| MAPEG | Membrane-Associated Proteins in Eicosanoid and Glutathione metabolism |
| NCIE | Nonbullous congenital ichthyosiform erythroderma |

| | |
|---|---|
| NICS | Neuro-immuno-cutaneous system |
| ORS | Outer root sheath |
| PADs | Peptidularginine deiminase |
| PALS | Papillon-Lefèvre syndrome |
| PBS | Phosphate buffered saline |
| PC | Pachyonychia congenita |
| PCR | Polymerase chain reaction |
| PKP | Plakophillin |
| PSEK | Progressive symmetric erythrokeratodermia |
| PSS | Peeling skin syndrome |
| RE | Restriction enzyme |
| RT-PCR | Reverse transcription PCR |
| SC | Stratum corneum |
| SG | Stratum granulosum |
| SLS | Sjogren-Larsson syndrome |
| SNPs | Short nucleotide polymorphisms |
| SPRs | Small proline-rich proteins |
| TEWL | Transepidermal water loss |
| TGase | Transglutaminase |
| TGN | Trans-Golgi network |
| TJ | Tight junction |
| TLC | Thin layer chromatography |
| *TMEM79* | Human transmembrane 79 gene |
| *Tmem79* | Murine transmembrane 79 gene |
| TBS | Tris buffered saline |
| TTBS | Tween-20 with Tris buffered saline |

# ABSTRACT

The skin barrier plays an essential role in protecting the body from external assaults and maintaining homeostasis. The differentiation of keratinocytes into toughened corneocytes begins in the basal layer of the epidermis and culminates in the formation of the stratum corneum (SC). Corneocytes are terminally differentiated cells which form a protective shield consisting of insoluble structural proteins cross-linked to the lipid envelope of the cornified cell envelope. In recent years, many molecular participants behind the extensive remodeling of the epidermis have been uncovered, and a spectrum of epidermal diseases have been shown to be associated with mutations coding for structural or enzymatic proteins involved in the complex processes of differentiation. However, genetic determinants for many epidermal barrier disorders are still unknown. The work in this thesis focuses on identifying the gene responsible for the phenotype of the *Matt*$^{ma/ma}$ *(matted)* mouse, which was first documented in 1957. To hunt this elusive gene down, extensive mapping of the murine epidermal differentiation complex (EDC) was carried out and the gene identified as transmembrane 79 *(Tmem79)*. Further work characterizing the human ortholog of TMEM79 showed that the protein localized to the membranes of lamellar granules (LG) in keratinocytes in the upper granular layers. Various human populations with skin diseases were screened, but no predisposition was identified between TMEM79 and allergic phenotypes.

In conclusion, *Tmem79* is responsible for the phenotype seen in the *Matt*$^{ma/ma}$ mouse, leading to spontaneous dermatitis due to a skin barrier defect. This novel gene encodes for a protein that is involved in desquamation of cornified layers, where an aberrant TMEM79 results in dysfunction LG transport/secretion.

# CHAPTER 1

# Introduction

# 1. INTRODUCTION

The differentiation of the epidermis is a concerted effort of maturation involving many players. However, there are still many questions regarding this sophisticated process. For instance, it is still unknown how the desmosome is modified to form corneodesmosomes, or how the nucleus and mitochondria are degraded during terminal differentiation. Further identification and understanding of the proteins involved in epidermal differentiation will not only aid us in understanding these processes, but will help to elucidate the mechanisms behind skin diseases and ultimately guide us in how to treat them.

## 1.1 STRUCTURE OF THE SKIN

The skin is the largest organ in the body and consists of one continuous surface. It provides protection by acting as a physical barrier against mechanical, chemical and thermal stress, as well as ultraviolet light. Small peripheral nerves within the dermis facilitate the ability to sense touch, pressure, pain and temperature. Thermoregulation of the body is regulated through body hair, sub-cutaneous fat, the action of perspiration and the vascular network under the surface of the skin through the increase or decrease in blood flow. Adipose fat tissue acts as a form of energy storage in the form of triglycerides, and the metabolic functions of the skin also include the synthesis of vitamin D.

Skin at different body sites varies in thickness, color and the presence of hairs and glands due to different functional demands. Despite these differences, there is a general structure to all skin types. The outermost layer of the skin is known as the epidermis, and it is thickest on palmoplantar skin. The epidermis lies above the dermis, which is nourished and connected to the rest of the body by the vascular

system. Epidermal appendages such as the hair follicles, sweat glands, sebaceous glands and nails extend upwards from the epithelium.

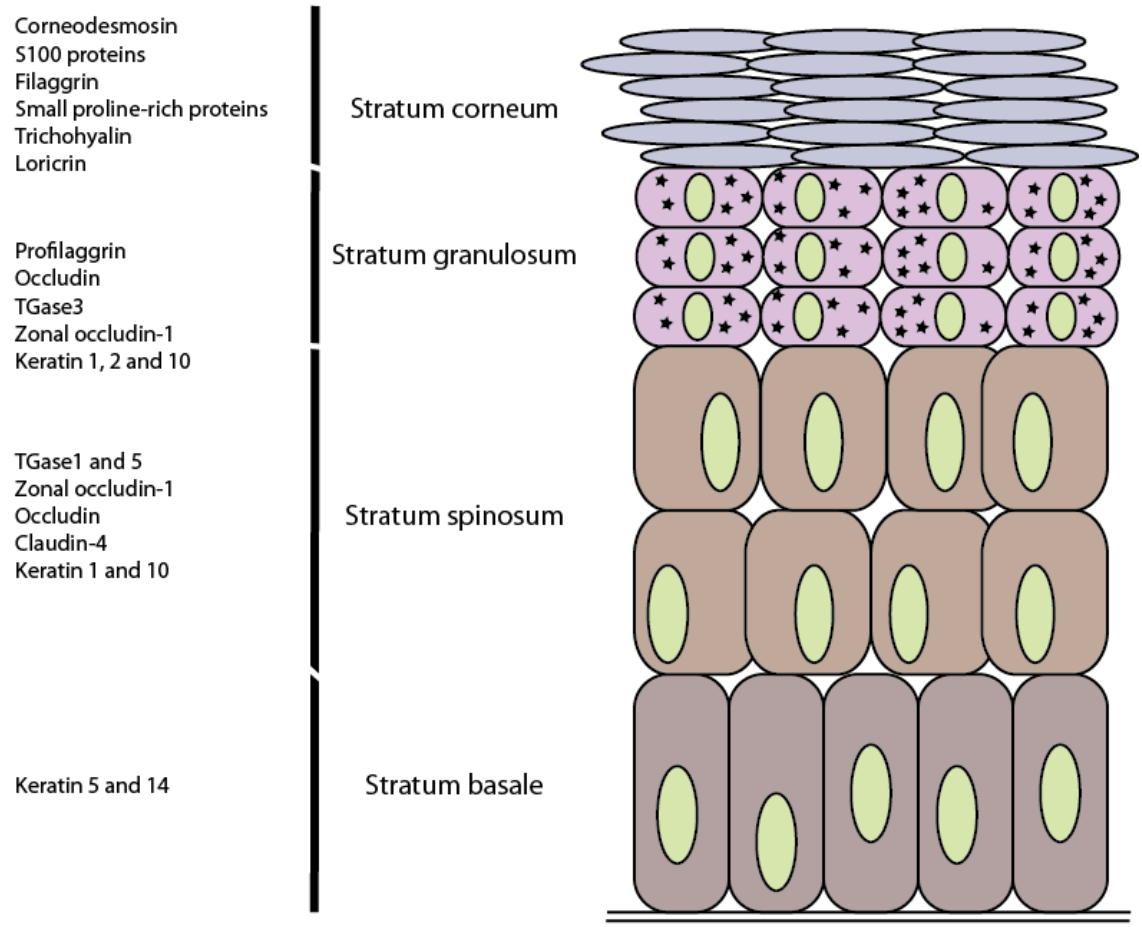



**Figure 1. 1 A cartoon of the epidermis with examples of layer-specific markers.** As keratinocytes differentiate and move upwards through the different layers, they express layer-specific markers.

Epidermal cells undergo growth and differentiation while migrating from the bottom to the top layer, acquiring distinct structural, functional and chemical characteristics in the process. Skin is characterized by squamous stratified cells with the ability for self-renewal and repair due to the presence of epidermal stem cells (ESCs). The differentiation of skin cells (herein called keratinocytes), are marked by the expression of various layer-specific markers (Figure 1. 1). The bottom-most layer is the stratum basale layer, consisting of a single layer of cuboidal cells that separate the epidermis from the dermis by a basement membrane. Basal cells are bound to the basement membrane by hemi-

desmosomes. ESCs are generally located within the stratum basale layer of interfollicular epidermis and also within the bulge region of the hair follicle. These cells divide asymmetrically, with one daughter cell undergoing differentiation and the remaining cell as the parent stem cell, designated to keep the population of keratinocytes constant within the epidermis (Blanpain et al., 2004). Cells within the stratum spinosum (spinous) layer are larger and polyhedral in shape. In the next layer, the stratum granulosum, granular keratinocytes are characterized by numerous keratohyalin granules (seen as star-shaped granules in Figure 1. 1) consisting of loricrin (L-granules) or filaggrin (F-granules) (Ishida-Yamamoto et al., 1993; Steven et al., 1990). It is within the granular layer that cells begin the final stages of terminal differentiation, secreting lamellar granules (also known as lamellar bodies or Odland bodies) into the intercellular space and crosslinking keratin filaments with the lipid envelope. Keratinocytes begin the process of degrading the nucleus and intracellular organelles in this layer so that when keratinocytes reach the stratum corneum, they are anucleate and devoid of organelles. This process is called cornification, where the cornified envelope is cross-linked to lamellae sheets within the intercellular space. The flattened keratinocytes of the stratum corneum (corneocytes) are shed at the surface during desquamation, a process which maintains skin homeostasis.

Skin pigmentation is conferred by the presence of melanocytes within the basal cell layer. The number of melanocytes are constant between individuals of different races and skin pigmentation is the result of the amount of melanin produced by the melanocytes (Burkitt et al., 1993; Iyengar, 2013). Melanocytes transfer melanin to neighbouring keratinocytes by using dendrites and this bestows protection of keratinocyte genetic material from UV radiation (Slominski et al., 2004).

The sensory component of the skin is known as the neuro-immuno-cutaneous system (NICS) and involves neurons, keratinocytes, melanocytes, Langerhans cells and Merkel cells. Merkel cells are located within the basal layer of the epidermis and are found in high concentrations in touch-sensitive areas. Under an electron microscope, they appear as a cluster of clear oval-shaped cells at nerve terminals (Halata et al., 2003). They act as mechanoreceptors and convert mechanical stimuli into nerve action potentials through the release of intracellular calcium and neurotransmitters (Slominski et al., 2004). Langerhans cells are present throughout the lower epidermal layers and have long dendritic processes extending between keratinocytes to detect pathogens before priming T cells within lymph nodes (Villadangos and Schnorrer, 2007).

There are two types of glands prevalent in the skin: the sebaceous gland and the sweat gland. The sebaceous gland is found throughout the human body apart from palmoplantar regions. They are usually found in the close vicinity of a hair follicle, forming the pilosebaceous unit. The main purpose of the sebaceous gland is to secrete sebum via the holocrine rupture (disintegration of whole cells) of mature sebocytes onto the surface of the skin (Smith and Thiboutot, 2008). There are glands on other parts of the body which carry out the same action; for example Fordyce spots on the lip and buccal mucosa, Meibomian glands and glands of the Zeiss on the eyelids and Montgomery areolar tubercles associated with lactiferous ducts. The exact function of sebum is unknown, although the overproduction of sebum has been identified as the major contributor of acne (Thiboutot, 2004).

There are two types of sweat glands in the human body; eccrine sweat glands found in most areas of the skin and apocrine sweat glands located in the axilla and genital regions (Burkitt et al., 1993; Zouboulis et al., 2007). Eccrine sweat glands are located on the majority of the skin surface, and are simple tubular glands that

secrete water and NaCl to the skin surface in response to increased temperature or fear-provoking stimuli. Apocrine sweat glands are associated with hair follicles and secrete a viscous and odiferous secretion into the hair follicles upon activation during puberty; e.g. pheromones.

## 1.2 THE SKIN AS A BARRIER

The skin is capable of self-renewal and its main purpose is to act as a barrier to protect the organism from the environment by preventing water loss and the entry of pathogens and allergens. There are two major components that make up the barrier: the stratum corneum and the lower nucleated epidermis. The cells of the stratified epidermis express specific cellular junctions, such as the desmosomes and adherens junctions which are mainly located within the basal, spinous and granular cell layers, and tight junctions and gap junctions within the granular layers. These intercellular junctions are critical in maintaining cell-cell adhesion, as evidenced by the wide range of genodermatoses associated with mutations in the genes that encode for juncitonal proteins, however, the focus of this thesis is on the role of the skin as a barrier, with particular emphasis on the formation of the stratum corneum.

### 1.2.1 Components of the stratum corneum

The physical protective functions of the skin are conferred largely by the stratum corneum. The cells of the stratum corneum consist of anucleated squamous epithelial cells, also known as corneocytes, embedded within a highly ordered non-polar lipid-enriched extracellular matrix. These lipids arise from the contents of lamellar granules (LG), and secretion occurs at the junction between the stratum

granulosum (SG) and stratum corneum (SC). The properties of the SC are made up of two parts: the protein envelope and the lipid envelope. The protein envelope consists of proteins covalently cross-linked by transglutaminases and sulfhydryl oxidases. In turn, the lipid envelope reinforces the protein envelope and provides the 'glue' that holds the cells of the SC together.

Transglutaminases (TGase) catalyze the covalent cross-linking between proteins via transamidation of glutamine residues in the presence of calcium (Hitomi, 2005). The highest levels of calcium are found within the granular layer, where calcium is important for both the regulation of transglutaminases and the initiation of epidermal differentiation (Hitomi, 2005). Expression of the various structural proteins of the SC is thought to be influenced by the presence of intracellular calcium. Within the skin, calcium is expressed at low levels in the basal level, lower levels in the spinous layers, the highest concentrations in the granular layer and low concentrations in the stratum corneum. Besides playing an important role in the initiation of keratinocyte differentiation and suppressing proliferation (Bikle and Pillai, 1993), it affects lamellar body secretion (Menon et al., 1994), and also proteolytic enzyme activity within the stratum corneum (Behne and Jensen, 2012). The impact of calcium signaling is induced by the interactions between calcium and calcium-sensing receptors (CaR) in keratinocytes and allows a direct feedback loop between the integrity of barrier function and a recovery response (Elias et al., 2002; Tu et al., 2001). Mutations affecting genes involved with the transport and secretion of calcium have been shown to lead to Darier's and Hailey-Hailey diseases, resulting in an intercellular split among suprabasal keratinocytes (Proksch et al., 2008).

Cells in the uppermost layer of the SG are linked to the SC by desmosomes, while adjacent corneocytes are connected by corneodesmosomes, both of which are vital during the process of desquamation (Proksch et al., 2008).

At the junctions of corneocytes, corneodesmosomes form intercellular covalent links with the cornified envelope. This junctional structure was identified as a marker for terminal epidermal differentiation, as they were exclusively identified in the SC (Serre et al., 1991) and will be discussed in Chapter 1.2.1.4.

#### 1.2.1.1 Proteins of the cornified envelope

The corneocyte is formed during the final stages of terminal differentiation where keratinocytes move towards the surface of the skin. During this process, keratinocytes lose their cellular organelles, become flattened and contain parallel bundles of keratin filaments. The cornified envelope is made up of a number of proteins which are cross-linked (Table 1. 1).

Involucrin (IVL) is located within the epidermal differentiation complex (EDC) and is encoded by one single exon. It has a high rate of polymorphism in humans and the number of potential glutamine residues available for cross-linking is thought to confer an evolutionary benefit (Eckert and Green, 1986). IVL makes up 2-5% of the protein in the cornified envelope (Nemes and Steinert, 1999). The expression of involucrin is initiated at the start of terminal differentiation in keratinocytes and is ubiquitously expressed in all skin sites (Robinson et al., 1997), and this is regulated via the distal or proximal-regulatory (DRR or PRR) elements found upstream of IVL. These two sites act in concert, and are thought to form a complex that interacts with basal transcription machinery to drive IVL transcription (Eckert et al., 2004).

| Molecules | Genes | Function | References |
|---|---|---|---|
| Involucrin | *IVL* | Structural protein of cornified cell envelope | (Candi et al., 2005) |
| Loricrin | *LOR* | Structural protein of cornified cell envelope | (Ishida-Yamamoto et al., 2000) |
| Envoplakin | *EVPL* | Structural protein of cornified cell envelope | (Kalinin et al., 2004; Karashima and Watt, 2002) |
| Periplakin | *PPL* | Structural protein of cornified cell envelope | |
| Small proline-rich proteins (SPRs) | *SPR1a* | Structural protein of cornified cell envelope | (Kartasova et al., 1996; Li et al., 2008) |
| Keratins | *KRT1, KRT10, KRT2* | Structural protein of intermediate filaments | (Nishifuji and Yoon, 2013) |
| Filaggrin | *FLG* | Aggregation of keratin intermediate filaments in cornified cells | (Sandilands et al., 2009) |
| Cystatin A | *CSTA* | Inhibitor of cell-cell adhesion | (Blaydon et al., 2011) |
| Corneodesmosin | *CDSN* | Component of corneodesmosomes | (Ishida-Yamamoto and Kishibe, 2011) |
| ATP-binding cassette subfamily A member | *ABCA12* | Lipid transporter responsible for cholesterol efflux from epidermal keratinocytes | (Sakai et al., 2007) |

**Table 1. 1 List of stratum corneum proteins involved in terminal differentiation of keratinocytes.** This process is tightly regulated and each protein has an important function in the maturation of keratinocytes into corneocytes.

Loricrin (LOR) is the most abundant protein in the epidermal cornified envelope. It is cross-linked into the cornified envelope at a late stage in differentiation and distributed in a diffuse manner across the cytoplasm of keratinocytes (Mehrel et al., 1990; Steven et al., 1990). LOR's glutamine-rich and lysine-rich domains are act as a substrate for TGase1 and TGase3 (Candi et al., 1995), and the N-terminal domain has been shown to interact with filaggrin and keratin 10 (Yoneda et al., 2012).

Small proline-rich proteins (SPRs) are a family of proteins with different expression levels in various epithelia, with SPR1a (cornifin α) and SPRR2 expressed in the epidermis (Fujimoto et al., 1993; Hohl et al., 1995). It is predicted that this gene family consists of two genes for SPRR1, seven genes for SPRR2, and a single gene each for SPRR3 and SPRR4 (Cabral et al., 2001; Gibbs et al., 1993). The proline-rich repeats are flanked by glutamine, lysine and proline residues, and similar to loricrin, are substrates for TGases. There appears to be a relationship between the amount of SPRs present in the cornified envelope and the epithelial sites expected to be exposed to trauma, for example, human buccal and gingiva contain >20% SPRs compared to human foreskin of ~5% SPRs (Steinert et al., 1998a; Steinert et al., 1998b). This group of proteins is thought to be expressed in response to an innate defense response after injury or after inflammation; for example, SPRR1B expression in response to inflammation at the ocular surface (Li et al., 2008), SPRR2A expression in response to gastric injury by *Helicobacter pylori* (Mueller et al., 2003) and SPRR4 epidermal expression in response to UV light (Cabral et al., 2001). An increase of SPRR1A expression in chronic lesional AD skin has been observed and is hypothesized to be a compensatory mechanism for the decrease of LOR (Jarzab et al., 2010).

Keratin intermediate filaments are integral components of the corneocyte, with keratin 1, keratin 2 and keratin 10 making up 80% of the protein content of a corneocyte (Eckert et al., 1997; Eckhart et al., 2013). Pairs of type I (acidic) and II (basic) keratin intermediate filaments make up the cytoskeleton, and type II keratin filaments are cross-linked via a single lysine residue to the cornified cell envelope during terminal differentiation in squamous cell epithelium (Candi et al., 1998b). This crosslinking allows integration of keratin intermediate filaments with the cornified cell envelope resulting in an insoluble corneocytic structure. Recently

it has been discovered that plasminogen activator inhibitor-2 (PAI-2) promotes differentiation by increasing expression of involucrin, keratin 10 and loricrin in cultured epidermal keratinocytes (Jang et al., 2010).

Profilaggrin (FLG) is encoded within the EDC and is the precursor for filaggrin, which is released by dephosphorylation and proteolysis at the N-terminus (Kam et al., 1993; Resing et al., 1993). After release from keratohyalin granules, filaggrin binds to keratin intermediate filaments causing them to form tight bundles that are a characteristic feature of the corneocyte. A proportion of filaggrin is cross-linked to cornified cell envelope proteins (Steinert and Marekov, 1995) and it is thought that the amino terminal fragments are incorporated into the cornified envelope (Dale et al., 1997). The process of deamination is a post-translational modification that occurs at the interface between the granular and cornified layer, and is catalyzed by the $Ca^{2+}$ dependent enzymes, peptidylarginine deiminase (PADs). Deamination is the conversion of arginine residues to citrulline residues and this process alters protein structure, increasing protein solubility and electrophoretic mobility and thus providing a better substrate for TGase crosslinking action. Keratin1, filaggrin and trichohyalin are all known substrates of PADs in the epidermis (Senshu et al., 1996; Tarcsa et al., 1996; Wang and Wang, 2013).

Cystatin A (*CSTA*) is a cysteine protease inhibitor that has been shown to be a potent inhibitor of house dust mite allergens Der p 1 and Der f 1 (Kato et al., 2005). Mutation of vital residues within *CSTA* has been predicted to alter protein tertiary structure and affect binding to potential cathepsin substrates, resulting in exfoliative ichthyosis. Within patients with loss-of-function mutations, there is loss in number and loosening of desmosomal junctions within the lower suprabasal layers as well as a decrease in the number of desmosomes (Blaydon et al., 2011).

In recent years, new late cornified envelope (LCE) proteins have been discovered (Henry et al., 2012). LCE 2B (formerly known as XP-5) is expressed late in envelope assembly and shares similarity with other cornified envelope proteins with regards to their amino acid composition (Marshall et al., 2001). Another LCE, NICE-1, was found to have the highest expression level in differentiating keratinocytes, and is now known as chromosomal 1 open reading frame 10 (C1orf10) (Marenholz et al., 2001). More recently, LCE3B and LCE3C have been associated with psoriasis, rheumatoid arthritis and allergic contact dermatitis (Ammar et al., 2012; Bergboer et al., 2012; Molin et al., 2011).

#### 1.2.1.2 Lipid envelope

Lipids extruded into the extracellular space hold the corneocytes of the SC together with corneodesmosomes. These lipids originate from lamellar granules (LG) by secretion from keratinocytes and contain acylated/glucosylated/hydroxylated ceramides, cholesterol and free fatty acids (Downing et al., 1993). LG are sensitive to calcium levels and during the ascent to the upper granular layers, deliver their contents into the intercellular space. These contents are further modified to form highly ordered lipid sheets that run parallel to the cell surfaces (Mandal and Downing, 1993). The LG membranes are absorbed into the plasma cell membrane and consist of lipids with sufficient length ($>C_{30}$) to span the lipid bilayer (of the plasma membrane) (Kalinin et al., 2001). The lipid envelope acts as a scaffold for the arrangement of intercellular lamellae. Following secretion of LG, polar lipids are enzymatically converted into non-polar lipids. The major lipids found in the stratum corneum are derived from the hydrolysis of glycosphingolipids into ceramides, conversion of phospholipids into free fatty acids and cholesterol which is derived from triglyceride catabolism (Downing et al., 1987). They play a vital

role in antimicrobial defense, by conferring an acidic pH to the stratum corneum, rendering it inhospitable to microorganisms (Behne et al., 2002).

Ceramide substrates become covalently attached to cornified envelope proteins such as involucrin (Nemes et al., 1999a; Nemes et al., 1999b), envoplakin and periplakin. Enzymes such as ceramide synthase (CerS) and ceramide hydroxylase/desaturase are involved in ceramide biosynthesis, and result in substrates such as are ω-hydroxyceramides (Marekov and Steinert, 1998; Mizutani et al., 2009). This accounts for 40% of the lipid content in the stratum corneum, but is absent in the granular, spinous and basal cell layers, and is characteristic of terminal differentiation. The role of esterified version of ceramide (ω-hydroxyceramide) is thought to act as a substitute for desmosomes in the stability of extracellular lamellar structures in cornified envelope cells (Wertz, 1997).

Free fatty acids and cholesterol are essential structural players in the formation of the skin barrier. The absence or reduced levels of either component results in a deficient skin barrier and can even result in lethality (Herrmann et al., 2003; Kelsell et al., 2005; Proksch et al., 1992; Thauvin-Robinet et al., 2005).

#### 1.2.1.3 Transglutaminases

Transglutaminases (TGase) 1, 3 and 5 are expressed in the epidermis. They play a vital role in the development of the cornified envelope.

TGase 1 is expressed in the granular layer of differentiating keratinocytes (Steinert et al., 1996a) and, after activation by proteolysis, is attached to the plasma membrane by its N-terminus (Steinert et al., 1996b). It plays several roles in cornified envelope formation: it aids protein envelope formation by cross-linking various structural proteins and in lipid envelope formation by cross-linking ceramides onto cornified envelope proteins (Nemes et al., 1999a; Nemes et al.,

1999b). TGase 3 has different substrates from TGase 1, and both have been shown to act on different lysine and glutamine residues of loricrin (Candi et al., 1995), and SPRs (Tarcsa et al., 1998). Expression of TGase3 is localized to the granular and cornified layers with a diffuse cytoplasmic distribution pattern. TGase 5 is also located within the granular layer and has been shown to cross-link loricrin, involucrin and SPRs (Candi et al., 2001).

#### 1.2.1.4 Corneodesmosomes

The discovery of a protein exclusive to the cornified layer, corneodesmosin (CDSN), revealed a marker for late epidermal differentiation. In the lower stratum granular layers, CDSN appears to be in the cytoplasm, but migrates to LG in the upper stratum granular layers, before finally heading to the center of corneodesmosomes in the SC (Serre et al., 1991). CDSN has a strikingly high serine and glycine content and is thought to confer adhesive properties at cell-cell junctions via its glycine loop domains (Steinert et al., 1991). Proteolysis of CDSN occurs in corneocytes committed to desquamation; a small degradation product of 33kDa is found in the outermost corneocytes, and the full-length form is absent (55kDa) (Serre et al., 1991). Besides being localized to the stratum corneum, expression of CDSN is also found within all three layers of the inner root sheath as well as the hair follicle (Gallinaro et al., 2004; Mils et al., 1992; Serre et al., 1991). It is vital in the desquamation process, where it is progressively proteolyzed by kallikrein-related peptidases (KLKs) and cathepsins (Bernard et al., 2003; Igarashi et al., 2004; Simon et al., 2001). As cells migrate upwards, desquamation of corneocytes occurs. Corneodesmosomes located on the apical and basolateral sides are degraded first, leaving those on the lateral sides intact up till the surface

of the skin (Haftek et al., 2011; Igawa et al., 2011). Skin diseases involved with abnormal degradation of corneodesmosomes are described in Chapter 1.4.1.

### 1.2.2 Components of the nucleated epidermis

Even though the stratum corneum performs the protective function of a physical barrier, the lower epidermal cell layers also play an important role as a second line of defense. The removal of the stratum corneum by a method called tape stripping is far from life threatening and results in an increase in transepidermal water loss (TEWL). Loss of the stratum corneum and granular layers through disease, e.g. staphylococcal scalded skin syndrome, is not life threatening (Elias et al., 1977). However, the removal of the suprabasal and subepidermal layers as a result of blistering diseases (e.g. pemphigus vulgaris, toxic epidermal necrolysis) and severe burns are life threatening due to extensive water loss and secondary bacterial infections. Foil or a grease ointment can be applied to act as an artificial barrier, reducing death rates significantly, demonstrating the importance of the lower epidermal layers (Honari, 2004).

#### 1.2.2.1 Tight junctions, adherens junctions and gap junctions

Tight junctions (TJ) form cell-cell junctions and act as a barrier, separating the apical and basolateral membrane and controlling the migration of paracellular molecules. Various TJ proteins have been identified in different layers of the epidermis. These include claudins 1, 3, 5, 7 and 17, occludin, JAM-A, zonal occludin-1, MUPP-1 and cingulin (Kirschner and Brandner, 2012). TJ barriers in keratinocytes are important in ion transport (Kirschner et al., 2009) and in preventing the entry of allergens and pathogens. It is thought that when Langerhans cells elongate their dendrites through the granular layer, the TJ barrier

is still intact (Kubo et al., 2009). Claudin-1 and 7 are found in all epidermal layers, occludin is localized to the granular layer and zonal occludin-1 and claudin-4 found in the suprabasal layers.

Adherens junctions (AJ) are situated directly below tight junctions, and are composed of a nectin-afadin complex and a cadherin-catenin complex. This junction provides the cell with the structurally adhesive connection with the actin cytoskeleton between cells. Nectin, α-catenin and β-catenin have been shown to bind to actin filaments (Drees et al., 2005; Niessen, 2007) and α-catenin is thought to regulate signaling receptors and affects asymmetric cell division in the basal layers (Lechler and Fuchs, 2005).

A major component of gap junctions are connexins (Cx), and these form hexagonal hemi-channels between adjacent cells to allow for the intercellular movement of signaling molecules or ions, or between the cytoplasm and extracellular space. Traffic of such molecules is dependent on activation by extracellular $Ca^{2+}$ concentrations that lead to conformational changes within the hemi-channel structure (Bao et al., 2004; Ebihara, 1996; Stout et al., 2002). Within keratinocytes, there is expression of Cx26 and Cx43 within human skin, but predominant expression of Cx43 in the basal and spinous layer of follicular and interfollicular epidermis (Salomon et al., 1994; Xu and Nicholson, 2013). Cx26 and Cx43 act together to regulate epidermal differentiation from basal keratinocytes, and was shown to severely impair overall growth and differentiation in in vitro experiments with rat epidermal keratinocytes (Langlois et al., 2007).

These junctions play an important role in preventing the entry of pathogens, for instance, *Clostridium difficile* secretes toxins that disassociate occludin, ZO-1 and ZO-2 at TJ of gastric epithelial (Nusrat et al., 2001), increasing epithelium permeability. Another method of entry is seen with the herpes simplex virus,

where the virus directly binds to nectin-1 of AJ in order to gain access into epithelial cells (Connolly et al., 2003). The invasion of *Shigella flexneri* via the gastrointestinal pathway occurs by manipulating Cx26 channels inducing a release of ATP resulting in the retraction of filopodia, allowing bacterial evasion (Romero et al., 2011).

#### 1.2.2.2 Desmosomes

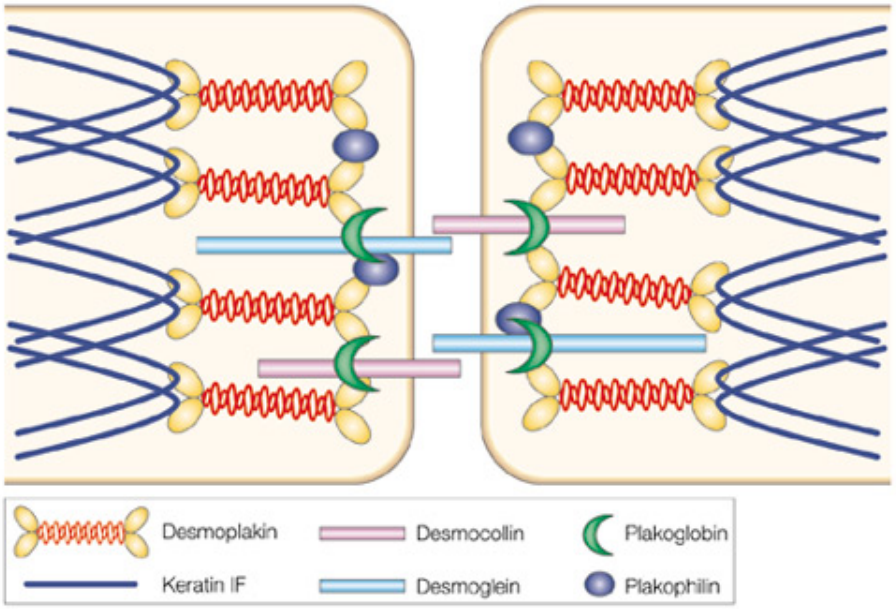



**Figure 1. 2 Simplified model of desmosomes.** Keratin intermediate filaments associate with the C-terminus of desmoplakin to confer intracellular structural stability to keratinocytes while desmoglein and desmoplakin form adhesion complexes for intercellular adhesion (Fuchs and Raghavan, 2002).

Desmosomes form the intercellular junctions between keratinocytes of the granular layer. Recently, a hypothesis suggested that desmosomes are "hyper-adhesive" structures. This is a term describing the state in which the strong adhesive state of demosomes was resistant to disintegrating in the absence of $Ca^{2+}$, unlike the other junctions (Garrod and Kimura, 2008; Hobbs and Green, 2012). They are formed from a complex of proteins that include desmoplakin, members of the armadillo family, desmosomal cadherins and accessory proteins essential for

adhesive function like Perp (P53 Apoptosis Effector Related to PMP22) (Ihrie and Attardi, 2005) and CDSN in modified desmosomes.

Desmoplakin plays an important role in desmosomal adhesion by interacting with keratin intermediate filaments through their carboxy termini (Figure 1. 2) (Godsel et al., 2008). Recent evidence also shows possible involvement in the regulation of microtubule organization (Lechler and Fuchs, 2007), as well as the association of desmoplakin with the actin cytoskeleton through their amino-terminal domains (Kalinin et al., 2002).Two structural homologues of desmoplakin, envoplakin and periplakin, have been shown to be located at the desmosomal junctions of epithelial cells (Ruhrberg et al., 1997).

Armadillo proteins include plakoglobin (PG; γ-catenin) and plakophillins (PKPs) 1-3. PG and PKPs show localization to the nucleus and adherens junctions, and are mostly found in desmosomes where they bind to DSGs and DSCs. All three members of the plakophillin (PKP) family are expressed in the epidermis with differentiation-specific patterns of expression (Garrod and Chidgey, 2008). This is mediated by p120-related desmosomal protein PKP1, which associates with actin and induces formation of filopodia (Hatzfeld et al., 2000).

Desmosomal cadherins include the three desmocollins (DSC 1-3) and four desmogleins (DSG 1-4). All of them are expressed in the epidermis in a differentiation-dependent manner (Garrod and Chidgey, 2008) and regulate intracellular signaling and differentiation through promoting the expression of molecules e.g. EGFR and MAPKs (Simpson et al., 2011).

Some of the diseases that occur when these intercellular junctions are compromised, affect tissues subjected to tremendous mechanical stress, such as the heart and skin. Mutations in PG result in arrhythmogenic right ventricular cardiomyopathy (ARVC), where patients are characterized by non-epidermolytic

PPK, wooly hair and heart failure. ARVC has also been linked to mutations in DSP, DSG2 and DSC isoforms (Al-Jassar et al., 2013).

Desmosomes mature into corneodesmosomes between the upper granular layers and the stratum corneum, and it is hypothesized that the desmosome is modified in some way (Kitajima, 2013). Studies of ultra thin sections show that the desmosome appears to have an electron-dense midline structure, and this structure is not seen in corneodesmosomes. Another difference is that in granular cells, keratin filaments extend into the desmosomal plaque, whereas in the corneodesmosomes, this association is no longer seen (Larabell et al., 1993).

### 1.2.3 Assembly of the cornified envelope

The initiation of terminal differentiation in keratinocytes is triggered by an increase in intracellular $Ca^{2+}$ concentration together with other signals that trigger the assembly of the epidermal cell envelope, summarized in Figure 1. 3 (Fischer et al., 2011; Kalinin et al., 2001)

Expression of periplakin and envoplakin occurs at the start of terminal differentiation. Periplakin is associated with keratin intermediate filaments and envoplakin is association with desmosomal junctions. Envoplakin and periplakin form hetero-tetramers that re-distribute to the plasma membrane in a $Ca^{2+}$-dependent manner. Shortly afterwards, TGase 1 and involucrin are expressed and associate with the membrane via its N-terminus (TGase 1) and glutamate residues (involucrin). Involucrin begins intra-chain head-to-head and head-to-tail crosslinking as well as inter-chain crosslinking with envoplakin and involucrin, to form an intracellular layer underlying the cell plasma membrane. This forms the scaffold, with the inclusion of desmosomes and other cell envelope proteins.

Simultaneously, lamellar granules (LG) fuse together with the cell membrane and secrete their contents into the intercellular space. The LG membrane, made up of long-chain ω-hydroxy ceramides become ester-linked to the scaffold proteins by TGase 1, eventually replacing the plasma membrane. The extruded lipids begin to form intercellular lamellae that appear as broad multilamellar lipid sheets (Landmann, 1986), and covalently attach to the long chain ceramides making up the lipid envelope in an interdigitating-fashion.

The reinforcement phase then begins, with TGase 3 cross-linking loricrin and SPRs together to form more insoluble complexes, which may act as substrates for subsequent crosslinking by TGase 1 to the scaffold (Steinert et al., 1999). The presence of SPRs increases the mechanical properties of the tissue, and it is thought that LEP/XP5 may act in a similar way. The keratin intermediate filament-filaggrin complex becomes cross-linked to the cornified cell envelope and lipid lamellae, forming the bulk of the corneocytes. There is bulk degradation of organelles such as the nucleus, mitochondria, ER and lysosomes, and although the mechanism is yet to be understood, there is evidence for the removal of nuclear DNA by DNase1L2 in the hair and nails (Fischer et al., 2011).

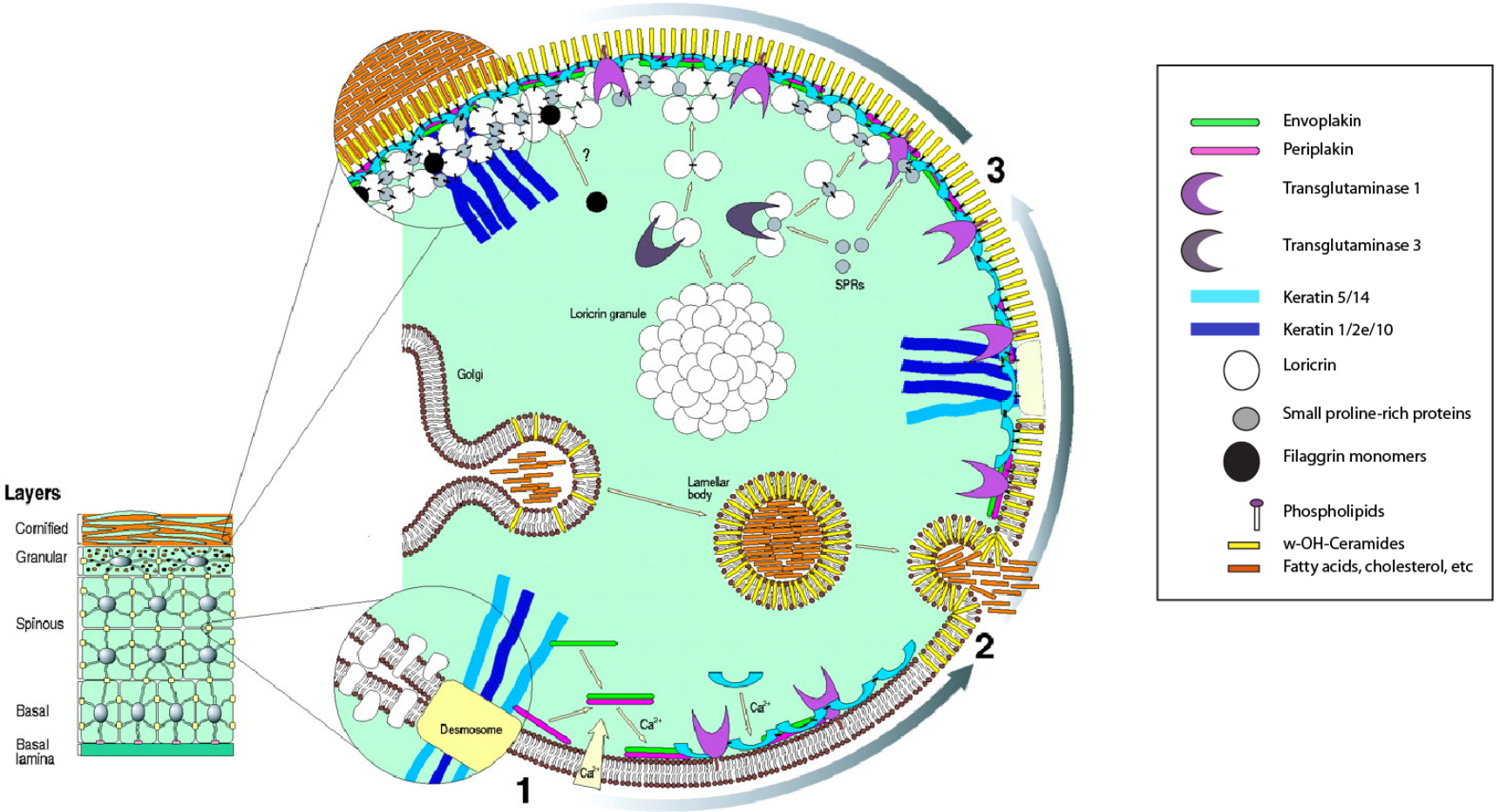



**Figure 1. 3 Assembly of the cornified cell envelope.** Begins in the spinous layers (1) with formation of a protein scaffold, followed by formation of the lipid envelope within the granular layers (2). In the final reinforcement phase (3), further crosslinking of CE proteins into the protein scaffold for structural stability in the cornified layers. Image adapted from Kalinin et al., 2001.

## 1.3 THE HAIR FOLLICLE

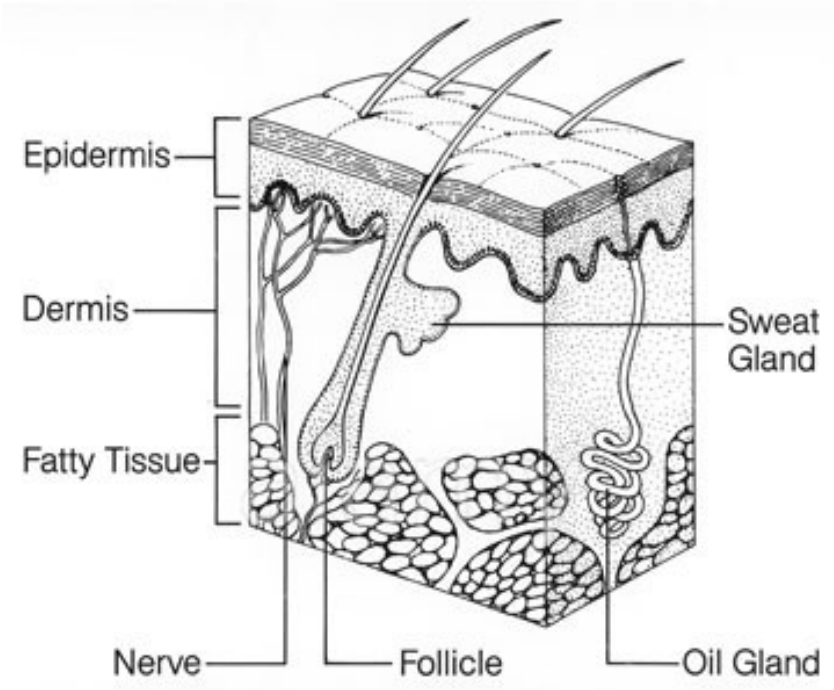



**Figure 1. 4 Diagram of the human hair follicle.** Other skin appendages are depicted within close proximity with the hair follicle. E.g. sweat glands, nerve endings and oil (sebaceous) glands. Stock image obtained from dijitalimaj.com.

Hairs are filamentous structures made of densely packed keratin filaments that project out from the surface of the skin. The lower follicle begins in the dermis, and the hair shaft extends out onto the surface of the epidermis (Figure 1. 4). Other appendages include the sebaceous gland and follicle pigmentation unit.

### 1.3.1 Structure

The human hair follicle is made up several concentric layers: radiating from the middle is the medulla, surrounded by the inner root sheath (IRS) followed by the outer root sheath (ORS) (Figure 1. 5).

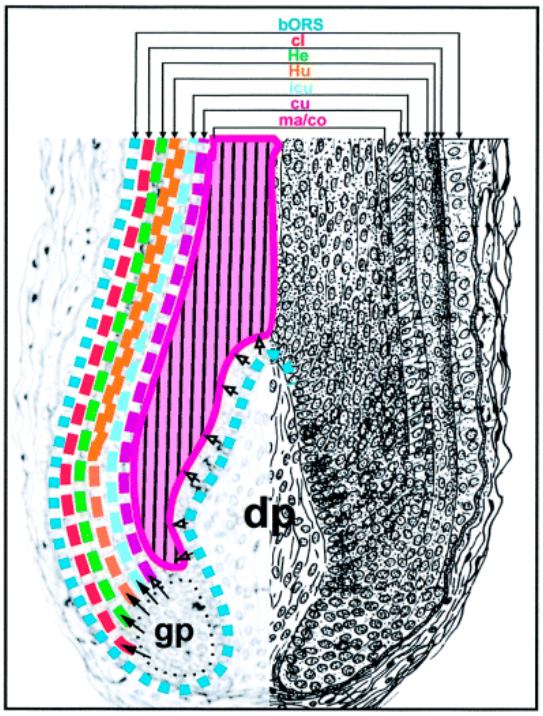



**Figure 1. 5 Cross-section of the human hair follicle.** The different layers are shown in different colours; bORS, basal layer of outer root sheath; cl, companion layer; He, Henle layer; Hu, Huxley layer, icu, inner cuticle layer; cu, cuticle layer; ma/co, matrix, cortex; dp, dermal papillae. Image taken from Langbein et al., 2002.

The epidermis is contiguous with the outer root sheath with two visible layers; the suprabasal layer expresses keratins 5, 6, 14, 16 and 17 whilst the basal layer expresses keratins 5, 14 and 19 (Winter et al., 1998). Cells of the companion layer (cl) lie between the ORS and IRS, and are morphologically distinct; they lack desmosomes and appear elongated and flattened, with expression of only one specific epithelial keratin, keratin 75 as well as non-specific keratins such as keratins 6, 16 and 17 (Langbein and Schweizer, 2005).

The inner root sheath links the hair follicle to the interfollicular epidermis and can be further divided into three layers starting from the innermost layer: the inner cuticle (icu), the Huxley layer (Hu) and the Henle layer (He). This transient compartment allows the exit of the hair shaft on the skin surface (Alibardi, 2004; Orwin, 1979) and eventually detaches from the hair shaft in the isthmus region, in the presence of sebaceous duct-secreted enzymes that degrade the IRS (Langbein

and Schweizer, 2005). Adherens junctions and desmosomes have been seen to connect cells between the IRS and the companion layer (Alibardi and Bernd, 2013; Orwin et al., 1973). The presence of such cell junctions allows the IRS to prevent water loss through the hair follicle, as well as preventing the entry of foreign pathogens into the follicle (Rogers, 2004). Henle cells can be identified by the appearance of trichohyalin in their cytoplasm and undergo keratinization (Orwin, 1971). The Henle layer expresses keratins 25, 27, 28 and 71, while the Huxley layer expresses the same keratins as well as a layer-specific keratin, keratin 74 (Langbein et al., 2010).

The cuticle layer (cu) encircles the hair-forming compartment and is made up of the cortex (co) and matrix (ma). The matrix contains proliferative cells surrounding the dermal papilla (DP). Synthesis and regulation of the hair follicle is dependent on various molecular factors and expression of receptors, leading to the deposition of keratin proteins within the cortex strengthened by disulphide bonds (Powell and Rogers, 1997; Wang et al., 2000). Several keratins have cuticle-specific expression: keratins 26, 72 and 73 (Langbein et al., 2010). Formation of inter- and intra-molecular disulfide bonds between keratin intermediate filaments and associated proteins gives rise to tightly packed cortex cells within the hair. Long-chain fatty acids have been observed in the scale cells of the cuticle, conferring properties of hydrophobicity to the hair surface (Lee, 2011). It is thought that these fatty acids are linked to the scale cell envelope by thioester bonds (Jones and Rivett, 1997). The matrix surrounds the DP, a group of cells responsible for follicle induction and hair growth.

### 1.3.2 Hair cycle

The hair cycle has classically been described as having three phases: anagen (fibre production/active growth), catagen (regression/breakdown), telogen (resting), before regrowth at the anagen phase again (Figure 1. 6). In humans, the length of time for each phase varies dramatically between follicles, from a few weeks to several years (Bernard, 2012), in a seemingly independent and random manner. Various signaling pathways and factors have recently been discovered to be are involved: Wnt, BMP, Shh, Notch and other signaling pathways involving EGFs and TGFα as well as other transcription factors (Lee and Tumbar, 2012). These different signaling cascades act in a concerted effort at various stages of hair follicle development (Baker and Murray, 2012).

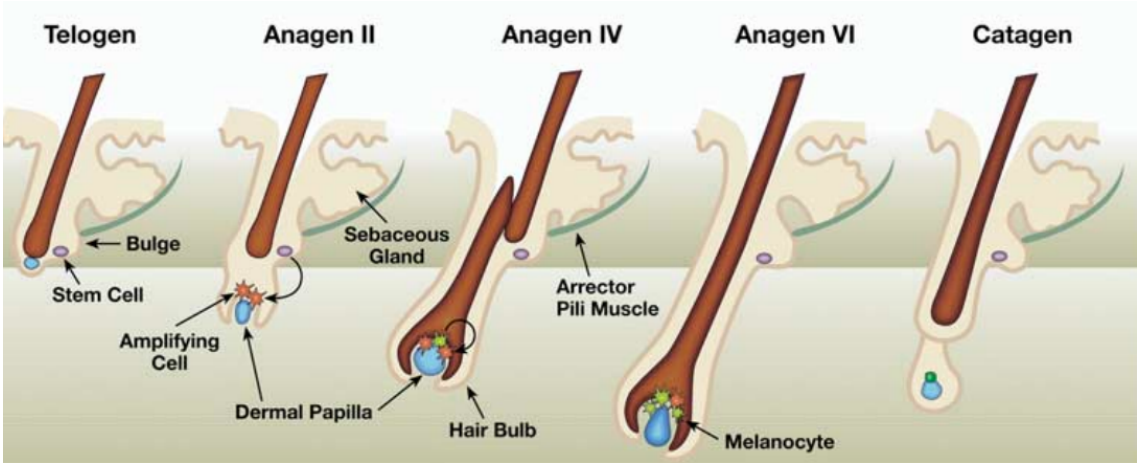



**Figure 1. 6 Diagram showing the three phases of the hair cycle**. During telogen, the hair follicle (HF) is resting, before entering and growing in anagen, and terminates differentiatin in catagen. IRS, inner root sheath; E, epidermis; HS, hair shaft; ORS, outer root sheath; DP, dermal papilla; M, matrix; CTS, connective tissue sheath; B, bulge; S, sebaceous gland (Steingrimsson et al., 2005).

During anagen, proliferating matrix cells move upwards and differentiate into one of the six lineages of the IRS and hair shaft (HS): Henley, Huxley, IRS cuticle layers and HS cuticle, cortex and medulla layers. Terminally differentiating HS cells extrude their organelles, resulting in a compact cell with bundles of cross-linked hair keratin filaments. It is suggested that the ORS is the result of a downward

migrating epithelium (Reynolds and Jahoda, 1991). Keratinization of the IRS gives it its ability to support the hair shaft during differentiation. The IRS-HS complex orientates itself by using the companion layer of the ORS as a slippage plane for movement towards the surface of the skin (Ito, 1986), and cells of the IRS release the HS as they reach the upper follicle. The coordinated expression of keratins is thought to be the result of locus control regions (LCR) that direct the activation of keratin genes by allowing physical accessibility of chromatin domains (Powell and Beltrame, 1994), resulting in a structure made up of Type I and Type II keratin intermediate filaments and a matrix of keratin-associated proteins (Powell and Rogers, 1997; Rogers, 2004). Between the IRS and HS, desmosomes, tight junctions and gap junctions are present in varying amounts. During keratinization, the junctions are replaced with a new cell membrane complex that develops into a continuous layer between the cells (Rogers, 1959). Wnt signaling is critical for the induction of hair follicle induction (Petiot et al., 2003). As the hair shaft increases in length and passes the lower third of the follicle, chromatin is degraded though the nuclear membrane remains insoluble and the keratinized cell forms an elongated structure in the cortex (Rice et al., 1994).

Between anagen to catagen, there is a temporary phase of amplification where matrix cells undergo a limited number of divisions before differentiating. Bmp signallng is critical for this transition (Andl et al., 2004), where the number of matrix cells decreases and HS and IRS differentiation is retarded. Various molecules have been identified in the anagen-catagen transition, but their exact mechanism and how they act together are still unknown. In various mouse models, a range of molecules have been implicated in the transition to catagen phase; FGF5, TGF-β1, BDNF (brain-derived neurotrophic factor), EGF, IGF-1 and BMPR1a (BMP receptor) (Andl et al., 2004; Foitzik et al., 2000; Hansen et al., 1997; Hebert et al.,

1994). Some factors that are known to retain the hair follicle in anagen phase are SGK3 and Msx2 (Alonso et al., 2005; Ma et al., 2003). Hair pigmentation also takes place during anagen, through regulation of the follicle pigmentation unit by melanocortins and a whole host of other factors (Kauser et al., 2005; Slominski et al., 2004).

At the start of catagen, there is a dramatic withdrawing of the hair follicle where apoptosis and terminal differentiation takes place (Lindner et al., 1997). As the HS stops differentiating, a structure called the club hair is formed at the bottom as it seals off. A temporary structure, called the epithelial strand, is formed between the receding lower follicle and the DP, which is destroyed as the DP moves upwards and rests beneath the remnant club hair. This regression appears to be linked to β-catenin, FGF5 and Dlx3 expression in the DP (Lee and Tumbar, 2012).

Hair follicles are in a dormant phase during telogen, and the growth is kick-started once again, when stem cells near the base of the follicle are activated to produce a new hair shaft (Blanpain et al., 2004).

## 1.4 DEFECTIVE SKIN BARRIER DISEASE

The cornified envelope (CE) is present in all stratified squamous epithelia and since its main role is to provide protection, any defects in the cornified layer are likely to result in disease. A number of genetic diseases have been discovered to be caused by mutations in genes that code for molecules involved with the cornified envelope. These defective molecules lead to changes in structural properties, metabolism of skin lipids or transglutaminase-related defects. Inherited skin disorders, or genodermatoses, fall into three categories; single gene disorders, chromosomal disorders and complex traits. In monogenic disorders, the inheritance pattern is normally autosomal dominant or recessive, with a smaller

number of X-linked dominant or X-linked recessive disorders. A pedigree may help to identify the pattern of inheritance. Some examples of autosomal dominant skin disorders are *KRT5* or *KRT14* in epidermolysis bullosa simplex (EBS), where dominant-negative mutant keratin proteins disrupts the ability of wild-type keratin to form intermediate filaments (Coulombe et al., 2009). In dominantly inherited skin disorders, haploinsufficiency may also be observed. This occurs when an affected individual has only a single functioning copy of a gene, resulting in an insufficient amount of functioning protein. For example, a mutation resulting in a prematurely truncating desmoplakin has been reported in individuals with striate palmoplantar keratoderma (PPK), where the dosage of desmoplakin is critical in maintaining desmosomal adhesion and structural integrity (Armstrong et al., 1999). In autosomal recessive disorders, two copies of the mutant allele are required in order for manifestation of the disease, these patients may be homozygous for the same mutation on each allele, or compound heterozygote for different mutations on each allele. Recessive diseases may be the result of consanguinity within the pedigree and are often severe. An example of a recessive autosomal disorder is harlequin ichthyosis (Chapter 1.4.2).

Not all disorders however, have complete penetrance, where all patients express the same clinical phenotype caused by the same mutant allele. Incomplete penetrance occurs when some individuals have the mutant allele but do not show evidence of disease. The variety of phenotypes seen in a single disease caused by the same mutation, could be caused by: modifier genes, which may alter the expression of another gene, environmental factors, allelic variation and complex interactions between genes and the environment (Lobo, 2008).

Genodermatoses result in a range of phenotypes, from dry skin to extremely severe forms that result in lethality, and these are described in greater detail in the

following sub-chapters. Although desmosomes do play a central role in causing skin barrier disorders (Al-Jassar et al., 2013), they have been briefly discussed in Chapter 1.2.2.2, and are not the focus of this chapter, which is centered on disorders relating to corneocytic components.

### 1.4.1 Skin disorders resulting from structural defects

The balance between proliferation and desquamation in healthy epidermis is a tightly controlled process, and many skin disorders are the result of altered structural proteins in causing a disturbance in the formation of the epidermis. Some examples are listed in Table 1. 2.

Point mutations within keratin genes have been shown to impair the assembly of keratin intermediate filaments (KIF), and therefore are mainly inherited in an autosomal-dominant manner (Porter and Lane, 2003). Pathogenic mutations tend to occur within the helix initiation and helix termination motifs of the rod-domain, which are critical for KIF assembly (Goldman et al., 1996; Haines and Lane, 2012; Uitto et al., 2007). Multiple keratin genes have been implicated in various skin disorders, and these are listed in Table 1. 2 (Szeverenyi et al., 2008). Mutations within the *KRT1* gene have been associated with epidermolytic hyperkeratosis (EHK) and epidermolytic palmar planter keratoderma (EPPK). These disorders are characterized by hyperkeratosis, skin fragility, a thickened stratum corneum and formation of blisters (DiGiovanna and Bale, 1994). The mutations in *KRT1* occur either within the rod-domain motifs as mentioned above, or within lysine residues on the head domain that are key for crosslinking KIFs to the cornified envelope (Candi et al., 1998b). In addition to causing disruption to the keratin filament cytoskeleton, mutations appear to affect lamellar granule secretion, resulting in an

abnormal cornified-envelope lipid scaffold (Schmuth et al., 2001). Several other keratins implicated in barrier disorders are listed in Table 1. 2.

| **Diseases** | **OMIM** | **Genes involved** | **Reference** |
|---|---|---|---|
| Epidermolytic hyperkeratosis (EHK) | #113800 | *KRT1* | (Rothnagel et al., 1992) |
| | | *KRT10* | |
| Epidermolytic palmoplantar keratoderma (EPPK) | #144200 | *KRT1* | (Covello et al., 1998) |
| | | *KRT9* | (Reis et al., 1994) |
| | | *KRT16* | (Terrinoni et al., 2000) |
| Ichthyosis bullosa of Siemens | #146800 | *KRT2* | (McLean et al., 1994) |
| Epidermolysis bullosa simplex (EBS) | #131900 | *KRT5* | (Dong et al., 1993) |
| | | *KRT14* | (Bonifas et al., 1992) |
| Pachyonychia congenita | #167200 | *KRT6A* | (Bowden et al., 1995) |
| | | *KRT16* | (McLean et al., 1995) |
| | #167210 | *KRT6B* | (Smith et al., 1998) |
| | | *KRT17* | (McLean et al., 1995) |
| | | *KRT6C* | (Wilson et al., 2010) |
| Vohwinkel syndrome (variant) | #604117 | *LOR* | (Maestrini et al., 1996) |
| Progressive symmetric erythrokeratoderma | #133200 | | (Ishida-Yamamoto et al., 1997) |
| Loricrin keratoderma | #604117 | | (Matsumoto et al., 2001) |
| Ichthyosis vulgaris (IV) | #146700 | *FLG* | (Smith et al., 2006) |
| Atopic dermatitis (AD) | #605803 | | (Palmer et al., 2006) |
| Peeling skin syndrome (PSS) | #270300 | *CDSN* | (Oji et al., 2010a) |
| Hypothrichosis simplex of scalp 1 | #146520 | *CDSN* | (Levy-Nissenbaum et al., 2003) |

**Table 1. 2 A list of skin disorders resulting from structural defects.**
Genodermatoses associated with defects in the structural components of the protein envelope within the cornified envelope and the papers in which they were first published (OMIM, Online Mendelian Inheritance in Man®).

The cornified envelope protein loricrin has been implicated in a variant form of Vohwinkel syndrome and progressive symmetric erythrokeratodermia (PSEK). An autosomal dominant single nucleotide deletion in *LOR* gives rise to a frameshift mutation, causing a loss of Gln and Lys crosslinking sites (Ishida-Yamamoto et al., 1997; Korge et al., 1997; Maestrini et al., 1996). A variant form of Vohwinkel's syndrome (VS) is characterized by palmoplantar hyperkeratosis and progressive constriction of the digits of the hands and feet, which may lead to auto-amputation. In classical VS, patients present with starfish-shaped hyperkeratosis and acoustic impairment caused by mutations in Cx26. Patients with PSEK also suffer from palmoplantar hyperkeratosis, along with erythematous plaques covering the body. The VS mutation identified by Maestrini *et al*. in a 4 generation pedigree, was a 1-bp insertion at the C-terminus (730insG) of *LOR*, the same mutation was reported by Matsumoto *et al.* in a Japanese family, described as having loricrin keratoderma. Both group of patients have very different clinical phenotypes. The Japanese family with loricrin keratoderma reported by Matsumoto *et al*., was initially diagnosed as having non-bullous congenital ichthyosiform erythroderma (NCIE), and were born as collodion babies with fine white scales in addition to constricting bands around the digits. They were identified as having a 1-bp insertion (709insC). In both cases, the insertions resulted in a replacement of the C-terminus with additional amino acids that would be expected to impair and disrupt cross-linking activity. Mutant loricrin has been shown to be retained within the nucleus due to the additional amino acids mimicking nuclear localization sequences (Korge et al., 1997). The reason for the difference in phenotype seen in these disorders could be because loricrin disorders are a heterogenous group of diseases. There does not appear to be any genotype-phenotype correlation, possibly because there were no proper clinical descriptors or because other mutations in other genes could be responsible

for the variant phenotypes. The variation in disease severity within families could also be linked to differences in expression or polymorphisms in the LOR gene (Ishida-Yamamoto, 2003).

Filaggrin (*FLG*) has been implicated in *ichthyosis vulgaris* (IV), which is a skin disorder characterized by dry, scaly skin and palmoplantar hyperlinearity with an increased risk of eczema, food allergies and asthma (McAleer and Irvine, 2013; Wells and Kerr, 1966). Patients have an absence or reduced levels of keratohyalin granules in the epidermis with mild hyperkeratosis. Several loss-of-function mutations have been discovered since the gene was first linked to IV (Hamada et al., 2008; Nomura et al., 2007; Sandilands et al., 2006; Sandilands et al., 2007; Smith et al., 2006). IV is inherited in a semi-dominant manner; heterozygous patients exhibit a milder phenotype compared to homozygous patients who have severe scaling and eczema (Smith et al., 2006). In these patients, a decrease or absence of filaggrin impairs skin barrier function. Further examination of structural abnormalities in IV patients revealed abnormal extracellular lamellar bilayer maturation and organization (Gruber et al., 2011). Along with a decrease of acidic filaggrin breakdown products, there is an increase in skin pH levels, which can result in the degradation of lipid-processing enzymes through the increased activity of skin serine proteases (Hachem et al., 2003). It has been shown that topical ceramide-dominant moisturizers help to replace deficient epidermal ceramides and to maintain the integrity of the skin (Sajic et al., 2012). *FLG* mutations R501X and 2282del4 were found to be strong predisposing factors for atopic dermatitis as well as other atopic diseases such as symptoms of allergies including asthma, food allergies and allergic rhinitis (Palmer et al., 2006).

Previously, nonsense mutations within the corneodesmosin (CDSN) gene were associated with hypotrichosis simplex of the scalp, an autosomal dominant disease

characterized by an initial loss of scalp hairs that progresses to complete baldness (Levy-Nissenbaum et al., 2003), but recently, nonsense mutations in CDSN have been linked with peeling skin syndrome (PSS) (Israeli et al., 2011; Oji et al., 2010a). This rare autosomal dominant disorder can be categorized into inflammatory or non-inflammatory with superficial peeling of the upper epidermis.

### 1.4.2 Skin disorders resulting from lipid-related defects

Impaired lipid synthesis may result in a group of disorders described as ichthyosiform disease, or in other words, cutaneous scaling. A decrease in the amount of epidermal lipids within the skin barrier would lead to dry skin, as the ability to prevent excessive loss of water is lost, through alteration of the epidermal lamellar granules (LG) (Jungersted et al., 2008; Khnykin et al., 2011; Menon et al., 1992). Some examples of lipid disorders are listed in Table 1. 3.

X-linked ichthyosis is an autosomal recessive disease caused by an accumulation of cholesterol sulfate within the extracellular space, due to a deficiency in steroid sulfatase (*STS*) (Basler et al., 1992). This inhibits the crosslinking activity of involucrin to the cell envelope by TGase 1 and esterification to the lipid envelope (Nemes et al., 2000).

A group of heterogenous disorders known as autosomal recessive congenital ichthyosis (ARCI) consists of lamellar ichthyosis (LI), nonbullous congenital ichthyosiform erythroderma (NCIE) and Harlequin ichthyosis (Oji et al., 2010b). The most severe congenital ichthyosis is the autosomal recessive Harlequin ichthyosis, where babies are born with a severe covering of thick white scales over their whole bodies, deep fissures in the dermis and severe out-turning of the eyelids and lips (ectropion and eclabium), renal dysfunction and cholestatic jaundice. Epidermal restrictions around the limbs restrict movement and the babies are at risk of auto-amputation of the digits. It was shown that mutations

within the ABCA12 gene (coding for adenine triphosphate-binding cassette transporter) were associated with this severe disease (Kelsell et al., 2005), with patients having nonsense mutations or frameshift mutations resulting in expression of truncated proteins (Akiyama et al., 2005; Thomas et al., 2006). The ABCA12 gene encodes a 6-transmembrane transporter protein that is localized to the LG in upper epidermal keratinocytes (Sakai et al., 2007). Patients suffering from Harlequin ichthyosis have absent extracellular lipid lamellar due to defective LG secretion (Dale and Kam, 1993). Mutations found within ABCA12 are mostly loss-of-function mutations, that are hypothesized to affect protein function by altering nucleotide-binding fold domains and/or transmembrane domains (Akiyama, 2010).

| **Diseases** | **OMIM** | **Genes involved** | **Reference** |
|---|---|---|---|
| X-linked ichthyosis | #308100 | *STS* | (Basler et al., 1992) |
| Harlequin ichthyosis (HI; ARC4B) | #242500 | *ABCA12* | (Kelsell et al., 2005) |
| Arthrogryposis, renal dysfunction and cholestasis (ARC) | #208085 | *VPS33B* | (Gissen et al., 2004) |
| Nonbullous congenital ichthyosiform erythroderma (NCIE) (ARCI2) | #242100 | *ALOX12B* | (Jobard et al., 2002) |
| | | *ALOXE3* | |
| Refsum disease | #266500 | *PHYH* | (Mihalik et al., 1997) |
| Sjogren-Larsson syndrome (SLS) | #270200 | *ALDH3A2* | (De Laurenzi et al., 1996) |
| Gaucher disease type II | #230900 | *GBA* | (Tsuji et al., 1987) |
| Cerebral dysgensis, neuropathy, ichthyosis, and plamaplantar keratoderma syndrome (CEDNIK) | #609528 | *SNAP29* | (Sprecher et al., 2005)<br>(Fuchs-Telem et al., 2011) |
| Chanarin-Dorfman syndrome (CDS) | #275630 | *CGI58* | (Lefevre et al., 2001) |

**Table 1. 3 A list of skin disorders resulting from epidermal lipid defects.** Disorders of the skin barrier caused by genes involved in the transport or metabolism of epidermal lipids within the cornified envelope and the papers in which they were first published (OMIM, Online Mendelian Inheritance in Man®).

Arthrogryposis, renal dysfunction and cholestasis (ARC) syndrome patients are thought to have an abnormal LG secretion and delayed desquamation, as they develop athrogryposis, renal-dysfunction, cholestatic jaundice and severe ichthyosis. ARC was linked to a loss-of-function mutation in VPS33B, a protein hypothesized to be involved in transportation and/or secretion of LG (Gissen et al., 2004). The VPS33 gene encodes a homologue of the yeast protein Vps33, which regulates vesicle-membrane fusion during late-stage protein trafficking (Huizing et al., 2001).

Reduced expression of SNAP29, a soluble n-ethylmaleimide sensitive factor attachment protein (SNAP) receptor (SNARE) protein, results in abnormal maturation and secretion of LG, causing a neuro-cutaneous disease known as cerebral dysgenesis-neuropathy-ichthyosis-keratoderma (CEDNIK) syndrome (Fuchs-Telem et al., 2011; Sprecher et al., 2005).

Nonbullous congenital ichthyosiform erythroderma (NCIE) shares similar characteristics with all the other lamellar ichthyosis (LI), but with fine white scales with nail dystrophy, scalp involvement and loss of eyebrows and lashes (Fischer et al., 2000). Mutations within lipoxygenase-3 or 12R are thought to affect the secretion of LG (Jobard et al., 2002).

Patients with Refsum disease have the main characteristics of retinitis pigmentosa, peripheral neuropathy, cerebellar ataxia and elevated CSF protein concentrations, along with minor ichthyosis, skeletal abnormalities and cardiac impairments (Skjeldal et al., 1987). This is the result of an accumulation of phytanic acid (long branched chain fatty acid) due to mutations within the phytanoyl-CoA hydroxylase (PHYH) gene (Jansen et al., 1997; Mihalik et al., 1997). An accumulation of phytanic acid is thought to interfere with peroxisomal-mediated import of proteins to organelles, although the exact mechanism is unknown (Waterham and Cregg,

1997). The symptoms of Refsum disease are later onset and have different characteristics than patients with Sjogren-Larsson syndrome (SLS), even though both disorders have been shown to accumulate phytanic acid.

SLS is characterized by ichthyosis, spastic diplegia or tetraplegia and mental retardation. Mutations in the fatty aldehyde dehydrogenase gene, ALDH3A2, affects the enzymatic oxidation of long-chain aliphatic fatty aldehydes to fatty acids, and have been discovered in patients (De Laurenzi et al., 1996). In addition to disrupted LG formation and barrier formation, the neurological symptoms associated with SLS may be due to the common ectodermal origin of the skin and brain, but more work needs to be done in order to identify the pathological mechanisms (Rizzo, 2013).

Gaucher disease is an autosomal recessive inherited disease where a subset of type II Gaucher patients present with early lethality by 2 years of age. Acute neuropathic type II form of Gaucher's has characteristics of hepatosplenomegaly, developmental regression, neurologic deterioration and cutaneous scaling abnormalities. Mutations in the β-glucocerebrosidase gene, an enzyme required for the hydrolysis of glucosylceramide into ceramide, impair the processing of lamellar lipids and affect the epidermal permeability barrier leading to an accumulation of glucosylceramide (Stone et al., 2000; Tsuji et al., 1987; Wigderson et al., 1989). The result of a ceramide deficiency in the contents of lamellar bodies leads to disordered lamellar-body sheets in the intercellular space of the cornified cells, with increased transepidermal water loss (Holleran et al., 1994). The defective β-glucocerebrosidase also disrupts macrophage or antigen-presenting cell functions, promoting inflammation and a loss of neuronal activity leading to CNS disease (Grabowski, 2012).

Chanarin-Dorfman syndrome (CDS) has similar features to NCIE, with fine white scaling, bilateral ectropion and eclabion and patients are often born as colloidion babies. In this autosomal recessive disease, there is an inherited lipid storage disorder caused by a mutation in the CGI-58 gene, resulting in an accumulation of lipid droplets in multiple tissues (Lefevre et al., 2001). It is hypothesized that mutations affect the interaction of CGI-58 protein with periphilin, which are unable to be localized to lipid droplets in adipocytes, resulting in abundance of abnormal LG and lamellar structures (Yamaguchi et al., 2004).

### 1.4.3 Skin disorders resulting from transglutaminase-related defects

Lamellar ichthyosis (LI) is an autosomal recessive life threatening disease that presents with characteristic dark brown, plate-like scales throughout the body (collodion membrane) at birth. Within the first few weeks of life, this membrane is shed, leaving the patient with scarring alopecia and a defective skin barrier (Huber et al., 1995b). It was discovered that several mutations within the TGase1 gene resulted in a decrease in enzymatic activity, these include nonsense, frameshift or splice-site alterations (Candi et al., 1998a; Huber et al., 1995a; Parmentier et al., 1995; Petit et al., 1997). The loss of enzymatic activity of TGase1 impairs the process of crosslinking of ceramide lipids to the CE, leading to barrier impairment. The location of TGase1 mutations appear to have an effect on disease pathophysiology, where patients with mutations p.Gly278Arg and p.Asp490Gly appeared to self-heal and the mechanism suggested is that these mutations lock the enzyme in an inactive *trans* confirmation under elevated hydro-static pressure (*in utero*), and when the water molecules are removed after birth, the enzyme is able to isomerized back to a partially active *cis* form (Raghunath et al., 2003). Other studies have shown that some LI patients have mutations in ALOXE3 and

ALOX12B, lipoxygenase genes on 17q13 (Krebsova et al., 2001; Osorio et al., 2012), and ABCA12 (ATP-binding-cassette transporter gene) (Lefevre et al., 2003) on 2q33-35 (Parmentier et al., 1996). Mutations within the ABCA12 causing LI are missense mutations, in comparison to HI, and predicted to have a less dramatic effect on protein function (Akiyama, 2010).

Recent work in patients with acral peeling skin syndrome (APSS, OMIM ♯609796) discovered prematurely truncating or missense mutations within the TGase5 gene (Cassidy et al., 2005; Kharfi et al., 2009; Pigors et al., 2012). Patients suffering from this autosomal recessive skin disorder present with blistering and peeling of the outermost layers of the palmoplantar epidermis (Shwayder et al., 1997). The absence of TGase5 activity severely impairs the incorporation of loricrin into the cornified envelope, increasing the amount of abnormal keratohyalin granules in the epidermal granular cells (Kalinin et al., 2001).

### 1.4.4 Skin disorders resulting from protease-related defects

Netherton syndrome (OMIM ♯256500) is caused by mutations in the SPINK5 gene, which encodes for a serine protease inhibitor, lymphyo-epithelial Kazal-type-related inhibitor (LEKTI1) (Chavanas et al., 2000; Sprecher et al., 2001). The loss-of-function mutations involved range from nonsense mutations, splice-site variants and insertion/deletions. This is an autosomal recessive disease where patients show congenital erythroderma, ichthyosis, trichorrexis invaginata and generalized atopy with high levels of IgE, hayfever and hypereosinophilia.

Mutations within cathepsin C gene (CTSC) have been linked to the Papillon-Lefèvre syndrome (PALS, OMIM ♯245000), where patients were characterized with palmoplantar hyperkeratosis and severe early onset of periodontitis (Hart et al., 2000). The exact mechanism is still undiscovered, although it is thought that

mutations within CTSC is associated with a reduced host response against infection, and is involved in epithelial differentiation or desquamation (Toomes et al., 1999).

## 1.5 MOUSE MODELS FOR SKIN BARRIER DISEASE

The mouse is well suited as an animal model for experimental use. They have the advantage of being small, easy and inexpensive to maintain, with a short lifespan. The mouse genome shares many similarities with humans in terms of genomic organization and is very well characterized. Inbred mouse strains make it easier for breeders to identify new phenotypes which can occur either spontaneously or through large-scale mutagenesis programmes.

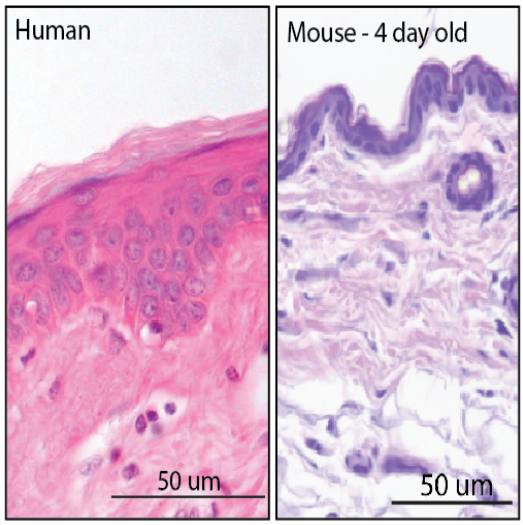



**Figure 1. 7 Comparison of human and mouse epidermis.** H&E stained sections of human skin and mouse dorsal skin. Mouse epidermis and dermis is significantly thinner compared to human epidermis, and comprises of more hair follicles.

The skin of mice is noticeably thinner compared to humans, with only three cell epidermal cell layers, whereas human epidermis usually comprises of 6 – 10 epidermal cell layers (Figure 1. 7). Skin from the footpad of the mouse most closely resembles human epidermis, as it is the thickest here. Ear epidermis is the thinnest, therefore this body site is typically used in sensitization studies in mice,

where the allergen is applied epicutaneously. Although the mouse is densely covered with hair and changes to the skin may not be immediately apparent, appendages such as the hair, whiskers, tail, and nails, are useful indicators when searching for mouse mutants with dermatological defects. For example, *flaky tail* mice which have a frameshift in the *Flg* gene, have pronounced tail constrictions. It is important to note that mutant and genetically engineered mice may not always replicate completely the same phenotype that is observed in humans. Many human skin diseases are not the result of a complete knockout of a gene, and merely have a reduced expression. An example of this is the ML.VS transgenic mouse model for Vohwinkel syndrome and PSEK, where a truncated form of loricrin (observed in patients) is expressed. The mutant form of loricrin appears to interfere with epidermal differentiation and homozygous knockout mice show a severe phenotype, but have a normal cornified envelope (Suga et al., 2000). These mice were observed to have severe erythroderma, epidermal barrier dysfunction, ichthyosis, tail constrictions and nonepidermolytic hyperkeratosis around the footpads. In loricrin knockout mice, there is increased susceptibility to mechanical stress, and newborn mice appeared to have erythroderma and shiny translucent skin (Koch et al., 2000). In another example, mice in which involucrin was completely knocked out by homologous recombination show a normal CE with no histopathological changes (Djian et al., 2000). Periplakin-null mice developed normally, and have a normal cornified envelope (Aho et al., 2004), as do envoplakin-null mice did not develop any substantial phenotype as well (Maatta et al., 2001). However, when a triple knockout mouse of involucrin, periplakin and envoplakin was made, the mice were observed to have delayed embryonic formation and postnatal hyperkeratosis, with structurally abnormal CE (Sevilla et

al., 2007). This suggests that there are compensatory effects by the other CE scaffold proteins in the single knockout mice.

Some mouse models do accurately reflect human disease and these include the Gaucher disease mouse model which was created by targeted ablation of the murine glucocerebrosidase gene (Tybulewicz et al., 1992). Homozygous mice die within 24 hours of birth and demonstrate accumulation of glucocerebrosidase in lysosomes, similar to the symptoms of the human disease. A *Spink5* deficient mouse with a complete absence of LEKTI presents with a similar features to Netherton syndrome, and studies showed that there was hyperactivity of stratum corneum tryptic enzyme (SCTE) and abnormal degradation of desmoglein 1 and desmoplakin resulting in a detached SC (Descargues et al., 2005). The complete absence of LEKTI was postnatally lethal, so the authors performed skin grafts from KO mice onto nude mice in order to observe phenotypic changes of the mutant mice. These grafts were observed to have acanthosis, hyperkeratosis of the SC, focal parakeratosis and hypergranulosis. Mice deficient for ABCA12 closely resemble the phenotype of Harlequin ichthyosis, with lamellar granule abnormalities and lipid envelope. These mice underwent rapid neonatal lethality, and showed a surprising involvement of ABCA12 in the lung, along with skin barrier formation (Yanagi et al., 2008). Alteration of other genes involved with lipid barrier formation proved lethal in the mouse, such as 12R lipoxygenase implicated in human lamellar ichthyosis and fatty acid transport protein 4 (FATP4) in ichthyosis prematurity syndrome (Nishifuji and Yoon, 2013).

In certain cases, it might be difficult to draw parallels between certain mouse models and human skin disorders as the mouse may not be responding in exactly the same way as in humans. Even though caution should be applied to data obtained from mouse models, they are still a good model organism for

understanding disease pathophysiology and molecular mechanisms in genodermatoses.

### 1.5.1 History of the *Matt*$^{ma/ma}$*Flg*$^{ft/ft}$ mouse

The *Matt*$^{ma/ma}$ *(matted)* mouse arose in 1952, in the University College London, as a spontaneous mutation in the CBA/Gr strain and was later outcrossed to the C57BL/Gr strain. By performing reciprocal backcrosses, it was determined that the gene responsible was a single fully penetrant recessive gene (Searle and Spearman, 1957). In the same paper, the authors went on to describe the phenotype of the *matted* mice. Homozygous mice had a distinctly fragile hair phenotype, with a propensity to develop alopecia. The coats were thin with clumps of hairs that were resistant to any sort of grooming. The loss of hair was also found to be exacerbated by scratching or by friction between the coat and the mouse's environment. Stress tests on the hair were carried out by bending individual hairs and this was found to cause splitting along the cortex, indicating reduced flexibility (Figure 1. 8a, b). The hair phenotype was more pronounced as the length of the hairs increased, which tended to split longitudinally due to brittleness. Originally, this phenotype gave rise to the initial notion that the *Matt*$^{ma/ma}$ mouse would be an appropriate animal model for the human disease, trichorrhexis nodosa due to perceived similarities between the hair defects of the human disease and *Matt*$^{ma/ma}$ mouse.

A second spontaneous mutation arose on the *Matt*$^{ma/ma}$ background in 1958; this mouse strain was identified as having a flaky tail (*ft*), and this double homozygote has been maintained at The Jackson Laboratory (Bar Habor, ME) ever since. The flaky tail phenotype is relatively mild, consisting of dry, flaky skin, with tail and paw constrictions, therefore the overt hair phenotype of the *Matt*$^{ma/ma}$ allowed the

*Flg*$^{ft/ft}$ mice to be identified. The matted hair phenotype appears earlier than the transient flaky-tail scaling phenotype and is permanently visible. The flaky tail mouse shows reduced profilaggrin expression and expresses a truncated version of the profilaggrin protein (~215kDa instead of >500kDa) (Presland et al., 2000). Involucrin and trichohyalin were expressed normally in these double mutants; therefore it was suspected that there was a mutation within the *Flg* gene, which was later identified as a frameshift mutation in the 6$^{th}$ filaggrin repeat (Fallon et al., 2009).

As a result of this work, the double mutant has been used as an animal model for studies of filaggrin deficient skin (Moniaga et al., 2010; Moniaga et al., 2013; Oyoshi et al., 2009; Scharschmidt et al., 2009).

The hairs from *matted* mice were examined microscopically (Jarrett and Spearman, 1957) and apart from having a defective cuticle with increased permeability to the medulla, other inter-follicular epidermal keratins were normal. The hair-cycle of the *matted* mice were normal for the first two generations, but had more rapid regeneration, possibly due to increased hair-loss in the animals owing to mechanical stress.

Subsequent genetic studies of the *matted* mice determined linkage of the *ma* gene with the histocompatibility loci on chromosome 3, with the locus an estimated 1.5 cM distance from the *ft* gene (Eicher and Lane, 1980; Lane, 1972; Lane and Eicher, 1979; Meisler and Seldin, 1991; Mobraaten et al., 1984)

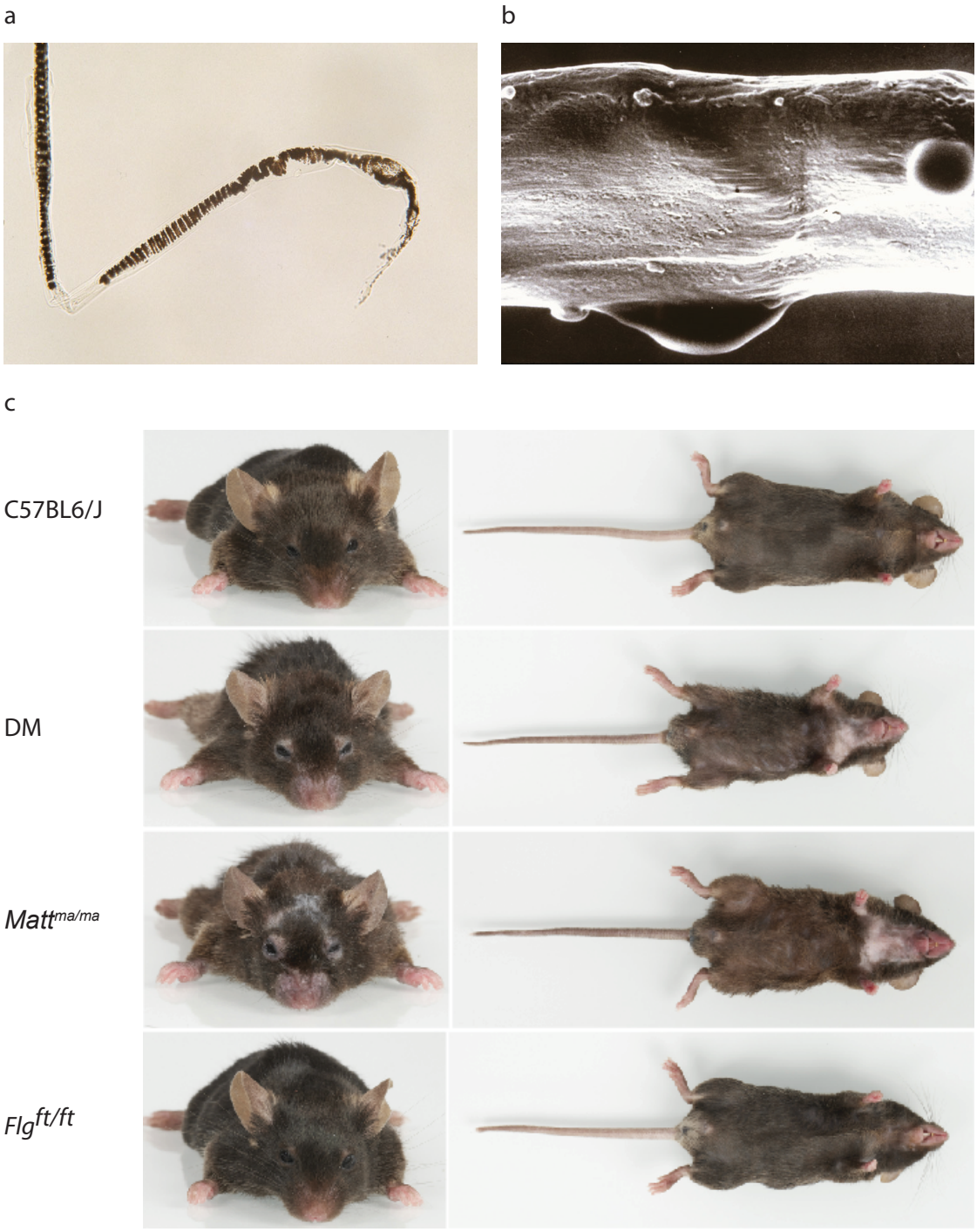



**Figure 1. 8 Phenotype of the *Matt*$^{ma/ma}$ *(matted)* mouse.** Hair-shaft fragility was demonstrated by stress tests (a) and hair phenotype was looked at in detail by scanning electron microscopy and showed rough hair cuticles (b; (Searle and Spearman, 1957)). Visually obvious differences between the wildtype C57BL/6J, *Matt*$^{ma/ma}$ *(matted)*, Flg$^{ft/ft}$ *(flaky tail)* and *Matt*$^{ma/ma}$*Flg*$^{ft/ft}$ double mutants (DM) as shown in panel. *Matt*$^{ma/ma}$ and DM mice show the characteristic matted phenotype and alopecia that was not seen in C57BL/6J or *Flg*$^{ft/ft}$ mice (c).

### 1.5.2 Recent work on the matted mouse

*Matt*$^{ma/ma}$ and double mutant (DM) mice develop significant skin inflammation with some exhibiting lesions and excoriation with blepharitis and periocular edema (Figure 1. 9). The development of progressive dermatitis-like skin inflammation in *Matt*$^{ma/ma}$ and DM mice was not seen in *Flg*$^{ft/ft}$. Our collaborators (Padraic Fallon et al, Trinity College, Dublin) looked at the immune profile of the various mouse strains in further detail and *Matt*$^{ma/ma}$ mice show elevated levels of IgE, IgG1 and IgG2a (Appendix I).

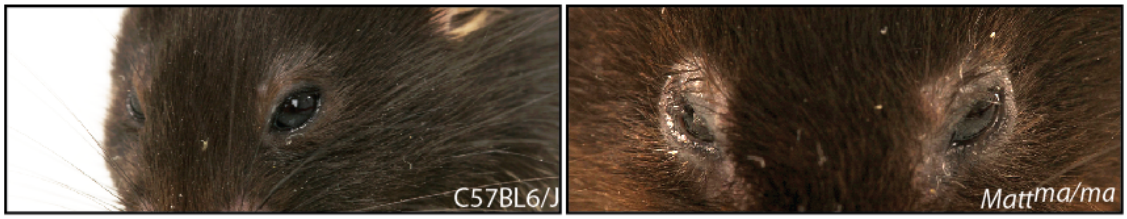



**Figure 1. 9 Significant ocular inflammation in the *Matt*$^{ma/ma}$ mouse.** Peri-ocular inflammation nvolving upper and lower eyelids, eyelashes and meibomian glands of the *Matt*$^{ma/ma}$ and DM mice that is not seen in the *Flg*$^{ft/ft}$ mouse.

Atopic dermatitis (AD) is a major inflammatory condition of the skin due to skin barrier deficiency, with mutations in *FLG* conferring incomplete penetrance to the development of AD. Although these DM mice develop spontaneous dermatitis, genetically engineered *Flg*$^{ft/ft}$ mice do not (Kawasaki et al., 2012). In this thesis, I set out to identify the gene responsible for the spontaneous dermatitis-like phenotype of the *Matt*$^{ma/ma}$ mouse, and to clarify its pathophysiological mechanisms.

# CHAPTER 2

# Discovery and identification of the *matted* gene

## 2.1. AIMS OF CHAPTER

Since the *Matt*$^{ma/ma}$ *(matted)* mouse represents a potential animal model for human atopic disease, it was necessary to identify the genetic mutation/s responsible for the *matted* phenotype. The goal of this chapter is to identify the causative gene and in order to achieve this aim, the regions of chromosomal crossover between *Matt*$^{ma/ma}$ and the backcross strain, C57BL/6J, had to be identified before candidate genes could be short-listed and investigated within this locus.

## 2.2. MATERIALS AND METHODS

### 2.2.1. The *matted* mouse strain

A genotyping method specific for the *ft* allele, Padraic Fallon and his group (Trinity College, Dublin) was used to separate the *ft* and *ma* alleles using a backcross-intercross protocol (Figure 2. 1, page 56). The matted hair phenotype allowed easy identification of the $Matt^{ma/ma}$ homozygotes and further backcrossing onto the C57BL/6J wild type strain. For simplification, the *matted* mice homozygotes will be referred to as $Matt^{ma/ma}$, flaky tail homozygotes as $Flg^{ft/ft}$, and double homozygotes as DM. All $Matt^{ma/ma}$ mice samples used in this chapter were F6 generation on the C57BL/6J background.

### 2.2.2. Mapping of homozygous and heterozygous regions

Skin samples were obtained from 4-day old neonatal C57BL/6J and $Matt^{ma/ma}$ mice (F6), and DNA was extracted using the DNA Purification Kit (Promega. Wisconsin, USA). Genomic DNA was amplified using GoTaq® Flexi DNA Polymerase (Promega, Wisconsin, USA). 50 ng of template DNA was amplified in a 25 µL reaction containing 1x PCR buffer (Promega, Wisconsin, USA), 1.5 mM $MgCl_2$, 5 pmol forward primer, 5 pmol reverse primer, 2.5 mM of each dNTP (Bioline), 4% DMSO (v/v) and 0.5 units of GoTaq® Flexi DNA Polymerase (Promega, Wisconsin, USA). PCR conditions used were: one cycle of 94°C for 5 min and 32 – 36 cycles of 94°C for 30 seconds, 54°C – 58°C for 30 seconds, 72°C for 1 minute with a final extension of 72°C for 5 minutes. All samples were sequenced using the ABI 3730 DNA System and was carried out by the DNA Sequencing Service (College of Life Sciences, University of Dundee). All mapping primers used to amplify and identify SNPs on mouse chromosome 3: 87,000,000 – 95,000,000 are listed in Appendix II. All mutations were confirmed by bidirectional sequencing. Matted genomic

sequences were compared to C57BL/6J to identify regions of homozygozity. This mapping strategy was employed initially, as previous genetic linkage studies showed that the *matted* gene was close to the *ft* gene, and initially seemed a fairly straightforward way of identifying the gene. The identification of a fairly large critical homozygous region prompted the use of next generation sequencing.

### 2.2.3. Next Generation Sequencing and bioinformatics

Mouse skin (30 mg) was disrupted using the TissueLyser LT (Qiagen Gmbh, Hilden, Germany) and total RNA was extracted using the RNeasy kit (Qiagen, Gmbh, Hilden, Germany) with the addition of Proteinase K (Roche, Mannheim, Germany) and RNase-free DNase (Qiagen, Gmbh, Hilden, Germany). The integrity and concentration of RNA was assessed by Agilent 2100 Bioanalyzer (Tayside Tissuebank, Dundee, UK). All samples had RNA Integrity Number (RIN) values between 8.60 – 8.90.

Approximately 10 μg of total RNA (three replicates for each genotype; C57BL/6J and $Matt^{ma/ma}$) was submitted for Next Generation sequencing to GenePool (University of Edinburgh, UK). Two replicates were run per lane on an Illumina GAIIx and HiSeq 2000 using v3 sequencing chemistry and the Roche 454 Titanium workflow that guaranteed 150 M reads per duplex. The raw fastq data provided had between 67 M – 105 M 76 bp paired-end reads per replicate and was analysed by Dr. Christian Cole (Bioinformatics Research Group, University of Dundee). Fastqc (v0.10.0 http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/) was used to check read quality before alignment to the Ensembl 66 mouse genome (Flicek et al., 2012) with the BWA short read aligner (Li and Durbin, 2009)(v0.5.9-r16 setting fastq format to Illumina v1.3+, parameter '-I')). Read counts were assigned to genes with HTSeq- count:

http://www.huber.embl.de/users/anders/HTSeq/doc/count.html using the following parameters: mode=union; type=exon; idattr=gene_id.

The Ensembl 66 GTF file was used for gene annotations, resulting in 28,007 genes expressed in at least one of the replicates. Differential gene expression between the C57BL/6J and the *Matt*$^{ma/ma}$ samples was performed with DESeq (v1.6.1) (Anders and Huber, 2010) and R (v2.14.0) with significant genes defined as those with $p<0.01$. In order to concentrate on genes with the most reliable read counts, genes with <120 reads in total (i.e. on average <20 reads per sample) across all six samples were removed, leaving 18,078 genes across the mouse genome.

### 2.2.4. SNP and InDel detection methodology

All replicates were combined into one BAM file per sample and the aligned reads were passed through SAMtools v0.1.18 (Li and Durbin, 2009) for calling SNPs and InDels with the standard command: samtools mpileup -C50 -DS –uf <ref> <bams>| bcftools view -bvcg > raw.bcf. The called SNPs/InDels were filtered to report only those with a minimum read depth of 20. SNPs/InDels were compared between the C57BL/6J and *Matt*$^{ma/ma}$ samples and only those found in the *Matt*$^{ma/ma}$ samples were retained for further analysis. 383 SNPs and InDels were identified in the EDC region (chromosome 3: 87,000,000-95,000,000) as being unique to the *Matt*$^{ma/ma}$ sample. The parameters and settings were determined by Dr. Christian Cole.

### 2.2.5. Analysis and identification of murine *Tmem79* gene

Each of the four exons was amplified individually by PCR using the genomic primers listed in Table 2. 1 using the following conditions for all exons: one cycle of 94°C for 5 minutes; 35 cycles of 94°C for 30 seconds, 54°C for 30 seconds and 72°C for 1 minute; and a final extension at 72°C for 5 minutes.

| Gene/mRNA | Location on mouse Chr3 | | F primer (5'-3') | R primer (5'-3') |
|---|---|---|---|---|
| *Tmem79* | 88132575: 88183555 | exon 1 | TTTGGAGACACAGGCAAG | TCTCTACAGCCAAGGGTCA |
| | | exon 2.1 | TTCCTCCCCACAGTGTTCT | AATGGCTCCTCCGACTTAG |
| | | exon 2.2 | CAGAGGCACCTACTAAGTCG | AGTGGCAGGTTCAGGATAC |
| | | exon 3 | GGTTTCTCTTCATTCTGCTG | AAGGCAATGGATTCAACAC |
| | | exon 4.1 | GGTGTTTGGAGTAGTATCGGC | TTTTCCTTCTCCTTGGCG |
| | | exon 4.2 | ACACTTCCTTCTTGGGCAG | GCTTGGAAGGCAGTTAGTTTC |
| | | exon 4.3 | CAAGCCTCTCTCTGTCTTCC | TGCTCCCAGTAAGTTGTCC |

**Table 2. 1** Primers used to amplify and sequence murine *Tmem79*.

### 2.2.6. Mouse genotyping assay

A 644 bp PCR product encompassing the third exon of *Tmem79* was digested with restriction enzyme *CviQI* (New England Biolabs© Inc., Massachusetts, USA) at 25°C for 1 hour and the digested fragments were separated according to size on a 1% (w/v) agarose gel by electrophoresis. Primers used for amplification of *Tmem79* exon 3 are shown in Table 2. 1. The $Flg^{ft}$ mutation was genotyped by Padraic Fallon (Trinity College, Dublin, Ireland) as previously described (Morita et al., 1995).

### 2.2.7. Semi-quantitative RT-PCR of *Tmem79* in skin

Mouse skin (30 mg) was disrupted using the TissueLyser LT (Qiagen Gmbh, Hilden, Germany) and total RNA was extracted using the RNeasy kit (Qiagen, Gmbh, Hilden, Germany) with the addition of Proteinase K (Roche, Mannheim, Germany). 1 µg of RNA was reverse transcribed to obtain cDNA using the ImProm-IITM

Reverse Transcription System (Promega, Madison, USA). PCR primers which spanned exons 3 and 4 of the *Tmem79* gene were used for amplification of cDNA. Primers to mouse Krt14 exon 8 were used as a positive control. Primers sequences are shown in Table 2. 2.

| Murine cDNA primers | F primer (5' - 3') | R primer (5' - 3') | Chromosome position | Expected mRNA size (bp) |
|---|---|---|---|---|
| *Tmem79* | CATCTTCTTCCC CTGTCTG | CAAGAGTGGCAG GAAAGTC | chr3:88,133,776-88,136,947 | 427 |
| Krt14 | ACTCACTCGCTC ACTTGCTCA | ATCTTGCTCTTC AGGTCCTC | chr11:100,068,255-100,068,804 | 580 |

**Table 2. 2** Primers used to amplify across exon 3 – 4 of murine *Tmem79* cDNA.

### 2.2.8. Quantitative PCR of *Tmem79* in multiple tissue array

Expression levels of murine *Tmem79* were measured across a multiple tissue panel which contained cDNA from 36 different NIH Swiss murine tissues (Origene Technologies, Rockville MD) using a TaqMan® Gene Expression Assay (Mm00470360_g1, Life Technologies Corporation, Applied Biosystems, USA). This probe spans exons 3 and 4 of *Tmem79* and gives an amplicon length of 89bp. Target location of the assay probe is Chr3: 88328653–88334433. Exons 3 – 4 of *Tmem79* were amplified with primers spanning exon 3 – 4 (Table 2. 2) and cloned into vector pCR2.1 (Life Technologies Corporation, Invitrogen, Maryland, USA) in order to provide a DNA standard for quantification of expression levels. Standard curve was calculated using the formulae provided by Applied Biosystems. ("http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_042486.pdf".)

cDNA samples extracted from *Matt*$^{ma/ma}$ and C57BL/6J epidermis were used as positive controls.

### 2.2.9. Immunofluorescent analysis of TMEM79 in mouse skin

Back skin was obtained from 4 day old *Matt*$^{ma/ma}$ and C57BL/6J neonatal mice and snap frozen in liquid nitrogen and stored at -80°C.

4-5 µm cryosections of mouse skin were cut and fixed in 1:1 methanol:acetone (-20°C), washed in PBS three times then blocked in 10 % (v/v) goat serum (Sigma-Aldrich, Missouri, USA) or 10 % (v/v) donkey serum (Sigma-Aldrich, Missouri, USA) for 20 minutes. Incubation with primary antibodies was carried out at room temperature for 1 hour. Sections were then washed three times with PBS, and subsequently incubated for 1 hour at room temperature with secondary antibodies. The same washing steps were repeated as above, and nuclei were counter-stained with 4,6- diamidino-2-phenylindole (DAPI), (Sigma-Aldrich, Missouri, USA) at a final concentration of 1 µg/mL. Sections were mounted in Prolong® Gold antifade reagent (Invitrogen, Life Technologies Corporation, Applied Biosystems, USA). Images were collected using either a Zeiss LSM700 confocal microscope or a Nikon ECLIPSE E600W microscope and analysed using SPOT imaging software v3.5.

The primary antibodies used and their working dilutions for immunofluorescence on tissue sections are listed in Table 3.3, page 79**.**

### 2.2.10. Nile red staining

A 0.5 µg/mL of Nile red working solution was freshly made up in a 75:25 glycerol-water mixture. Frozen skin sections were incubated for 5 mins with Nile red working solution and 1 µg/mL of DAPI (Sigma-Aldrich, Missouri, USA). Sections were mounted immediately and imaged.

### 2.2.11. Lipid analysis

Lipid analysis of *Matt*$^{ma/ma}$ and C57BL/6J whole skin was outsourced to My1nefield Research Services (Dundee, UK). Lipids were extracted from frozen skin using an adaptation of lipid extraction from animal tissues using chloroform-methanol-potassium chloride in the ratio 2:1:0.88 v/v/v (Folch et al., 1957). Three types of analysis were carried out:

1-D lipid class identification – murine skin extract was dissolved in solvent and spotted onto a thin liquid chromatography (TLC) plate together with known standards and separated in one direction using 80:20:2 iso-hexane/diethyl ether/formic acid.

2-D phospholipid species identification – phospholipid species were separated from murine skin extract using two solvent systems; chloroform/methanol/water (65:25:4) followed by chloroform/methanol/acetic acid/water (80:12:15:4).

1-D separation for ceramides – Murine skin extract was separated in one directiom using three different solvent mixtures; dichlormethane/ethyl acetate/acetone (80:16:4), chloroform/acetone/methanol (76:8:16) and isohexane/chloroform/acetone (6:80:14).

## 2.3. RESULTS

### 2.3.1. Identification of retained *ma/ma* regions

In order to identify the region harbouring the *matted* gene, a breeding strategy for maintaining a colony of *Matt*$^{ma/ma}$ mice was devised by our collaborators in Trinity College, Dublin. The matted hair phenotype allowed easy identification of *Matt*$^{ma/ma}$ homozygotes and therefore, the matted mutation was selected and retained (Figure 2. 1).

The *Matt*$^{ma/ma}$ mouse line underwent successive backcrosses onto the C57BL/6J background. Figure 2. 2 shows the breeding strategy that was used. As the successive rounds of backcrosses between C57BL/6J wild type mice and matted mice proceeded, the disease-causing locus would be expected to decrease in size due to random meiotic recombination, resulting in two uneven flanking regions of heterozygosity. We had also previously sequenced a BAC clone prepared from *Matt*$^{ma/ma}$*Flg*$^{ft/ft}$ double homozygotes, which effectively excluded Hrnr and Flg-2 as the causative gene (Dr Aileen Sandilands, University of Dundee). The mapping experiments carried out in this chapter, were on C57BL/6J and *Matt*$^{ma/ma}$ F6 generation progeny, with mapping primers designed across the region of chr3: 80,000,000-95,000,000 (Appendix II) to identify any genomic differences between C57BL/6J and *Matt*$^{ma/ma}$ mice. The primers were designed at regular intervals with special attention paid to cover non-annotated genes. Approximately 1000 mapping primers were designed and a total of 142 variants between C57BL/6J and *Matt*$^{ma/ma}$ were identified. All sequencing reads were mapped against the reference mouse genome (July 2007 (NCBI37/mm9)) and two regions of heterozygous regions (blue) were identified which flanked the homozygous region (red) that contained the mutation-harboring gene (Figure 2. 3a). A complete list of all the variants identified are listed in Appendix III, with a selection represented in Figure

2. 3a. The homozygous region was narrowed to an area between chr3: 87,336,555-88,392,125(Figure 2. 3b).

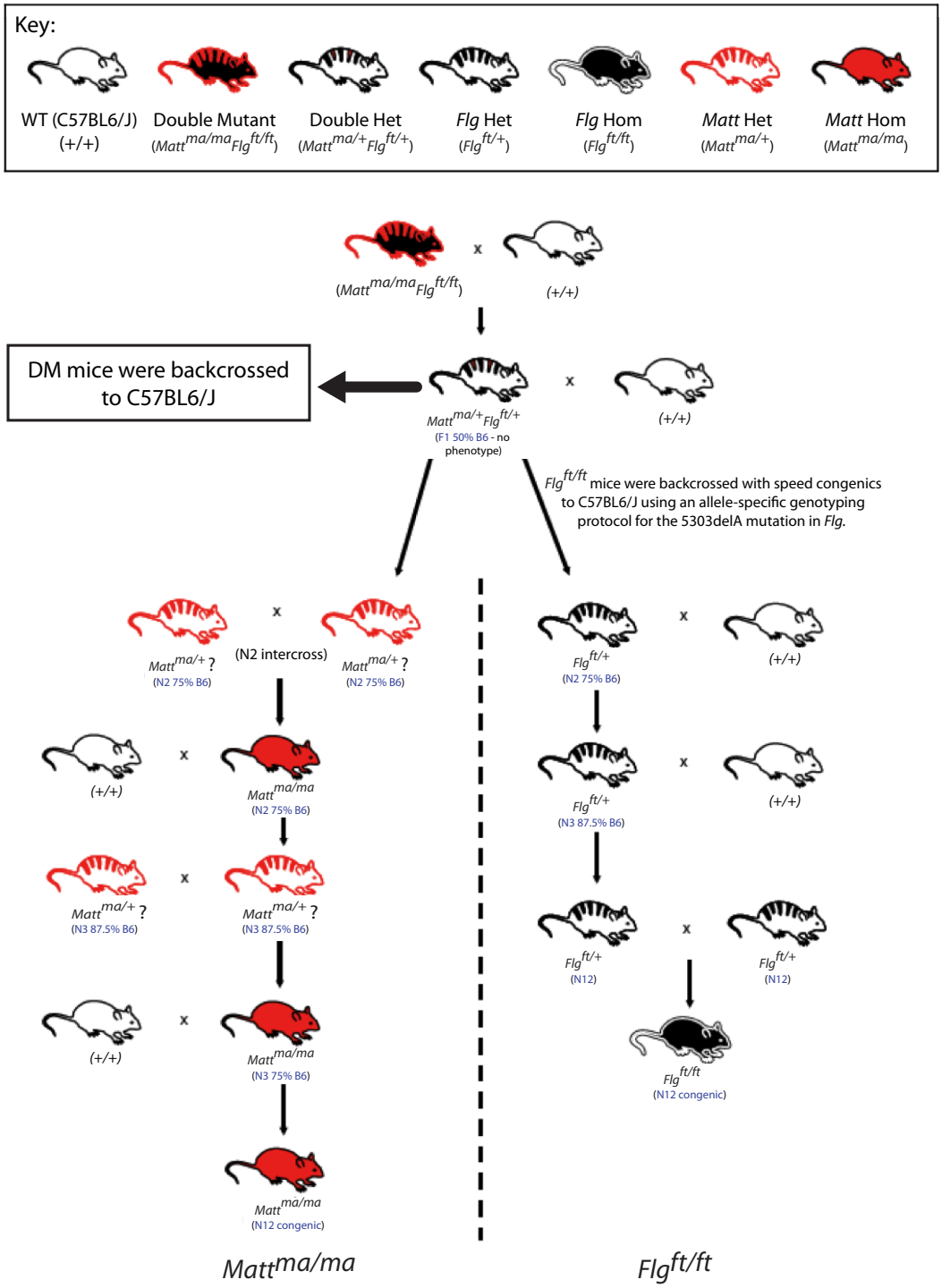



**Figure 2. 1 Breeding strategy employed to isolate the *Matt*$^{ma/ma}$ mouse.** A series of reciprocal backcrosses and intercrosses were performed to isolate *Matt*$^{ma/ma}$ from the *Matt*$^{ma/ma}$*Flg*$^{ft/ft}$ double mutant by our collaborators from Trinity College, Ireland. % refers to congenicity of mutant mouse strains with C57BL6/J wildtype mice.

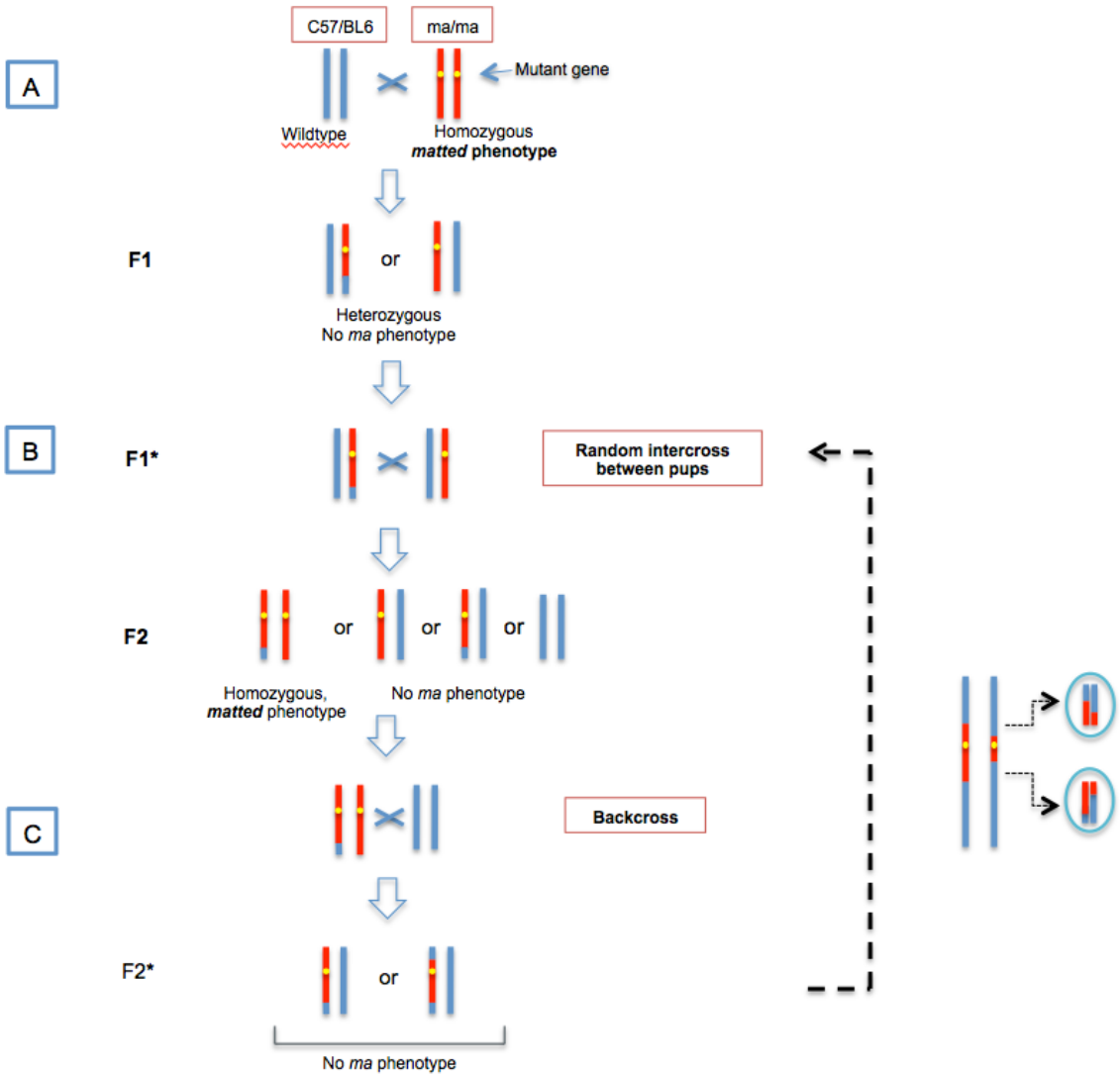



**Figure 2. 2 Schematic cartoon demonstrating the hypothesis behind reducing the critical *matted* homozygous region**. Backcrosses between wildtype (C57BL/6J, blue chromosomal regions) and *Matt*$^{ma/ma}$ (red chromosomal regions) harboring the mutant *matted* gene mutation (yellow dot) to give rise to heterozygous offspring (F1). Random intercrosses between F1 offspring result in homozygous *Matt*$^{ma/ma}$ F2, which are selected by their phenotype. Backcrossing between homozygous *Matt*$^{ma/ma}$ F2 and wildtype is repeated to obtain a congenic *Matt*$^{ma/ma}$ strain with C57BL/6J background. Each backcross results in random meiotic recombination and as a result, the region of *Matt*$^{ma/ma}$ surrounding the matted gene is reduced.

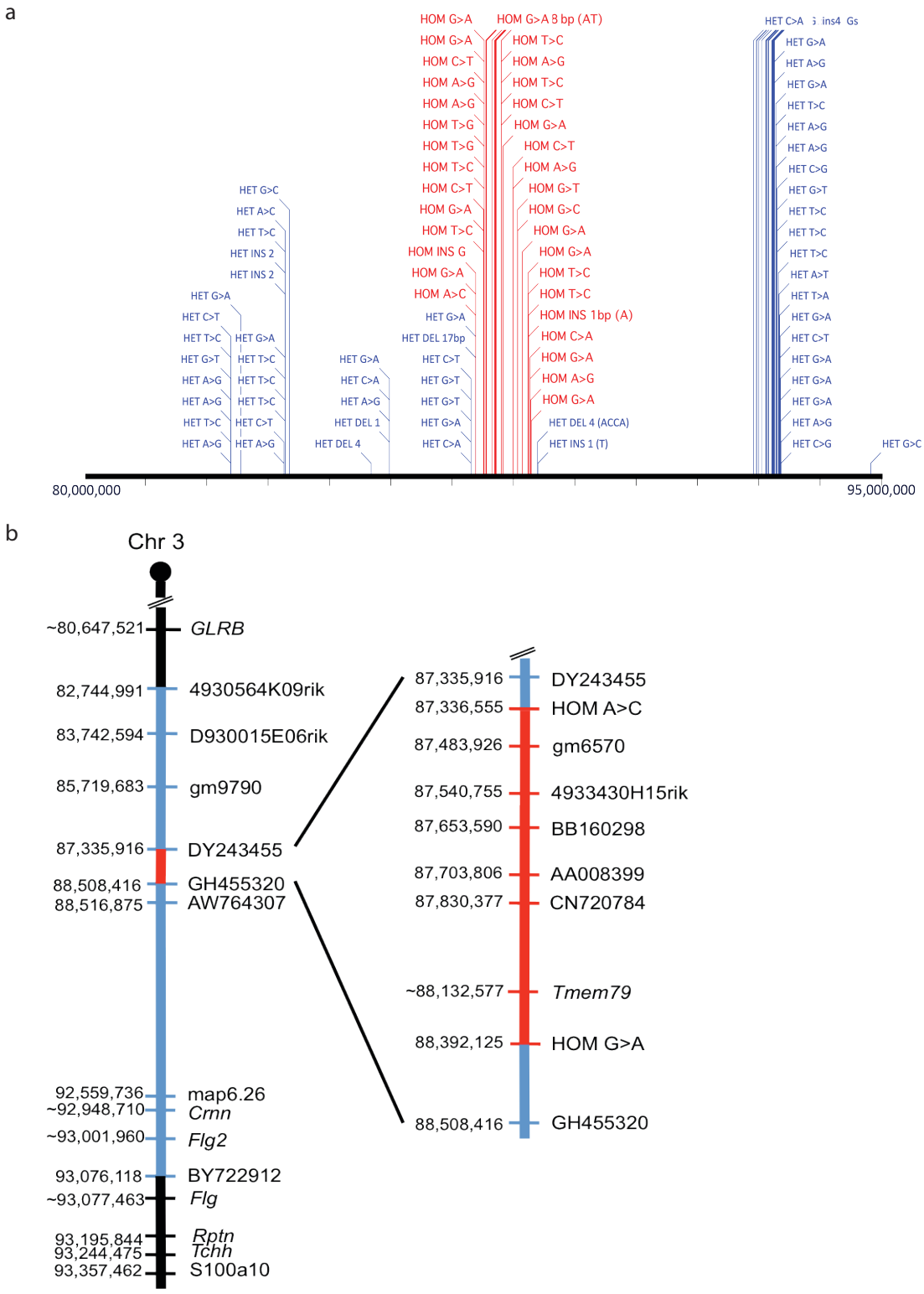
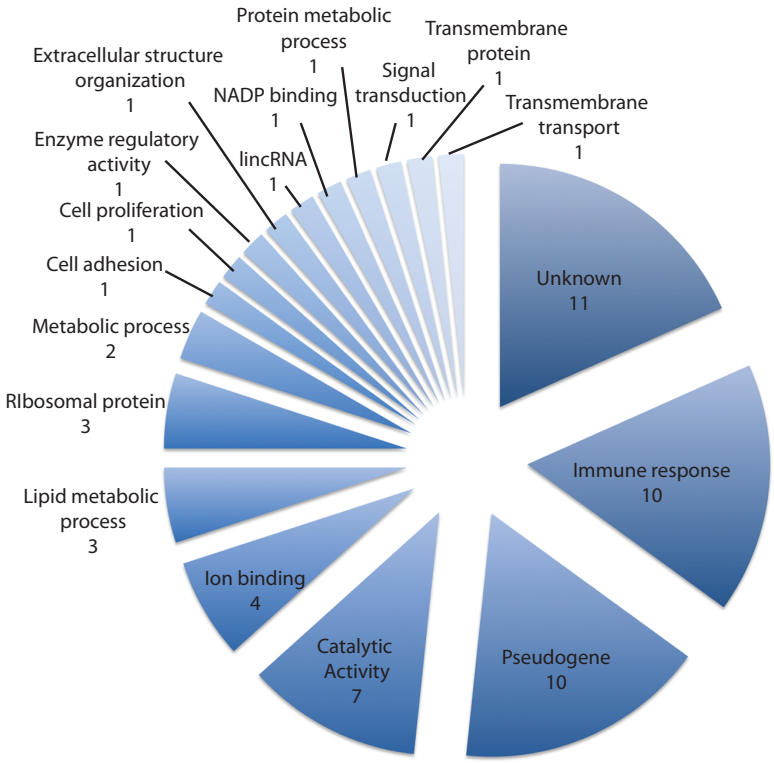



**Figure 2. 3 Representation of mutations discovered within the critical homozygous region of interest.** Variants identified with murine chr3: 80,000,000 – 95,000,000 were mapped using MacVector v12.7 (a). An overview of heterozygous regions flanking the retained homozygous *Matt*$^{ma/ma}$ area containing the mutation-harboring gene on chromosome 3 and the distance from the *Flg* gene (b). For both panels, the heterozygous regions are denoted in blue and retained homozygous *Matt*$^{ma/ma}$ regions in red.

### 2.3.2. Next Generation Sequencing and bioinformatics

Reads obtained from the transcriptome sequencing of C57BL/6J and *Matt*$^{ma/ma}$ skin RNA were compared to each other and a total of 383 SNPs and InDels were identified within the epidermal differentiation complex on chromosome 3 (region between chr3: 87,000,000 – 95,000,000; Appendix IV). Based on the information obtained by genomic mapping, the critical region of homozygosity between chr3: 87,336,555 - 88,392,125 yielded a shortlist of 210 SNPs within potential candidate genes. These potential candidate genes were analyzed by locating all possible open reading frames (ORFs) for potential protein coding regions, and determining whether the mutations introduced had any impact on protein coding. This was achieved using a software program called the Integrated Genome Browser (Nicol et al., 2009). Several mutations that affected coding were uncovered within potential candidate genes, but a single nonsense mutation c.840C>G (p.Y280*), within the *Tmem79* gene was identified within the homozygous 1.05 Mb interval. Comparison of gene expression levels between the *Matt*$^{ma/ma}$ and C57BL/6J mice identified 114 genes with differential expression (Appendix V). There were 60 up-regulated genes and 52 down-regulated genes across the genome of the *Matt*$^{ma/ma}$ mice using the most up-to-date version of DSeq with $p < 0.01$ (March 2012). They were categorized according to their Gene Ontology description (Figure 2. 4). The largest group of up-regulated genes was related to immune response, which would correlate with the observed atopic phenotype of *Matt*$^{ma/ma}$ mice. *Tmem79* was under-expressed in *Matt*$^{ma/ma}$ mice (base mean of 695), with a differential expression of -1.75 log2fold change in comparison to C57BL/6J (base mean of 2330). The other genes of interest encoded for proteins with functions relating to “catalytic activity”, ion binding and metabolic processes involving fatty acid and lipids, with the largest group of genes having an unknown function.

a Upregulated genes in the *Matt*$^{ma/ma}$ mouse




b Downregulated genes in the *Matt*$^{ma/ma}$ mouse

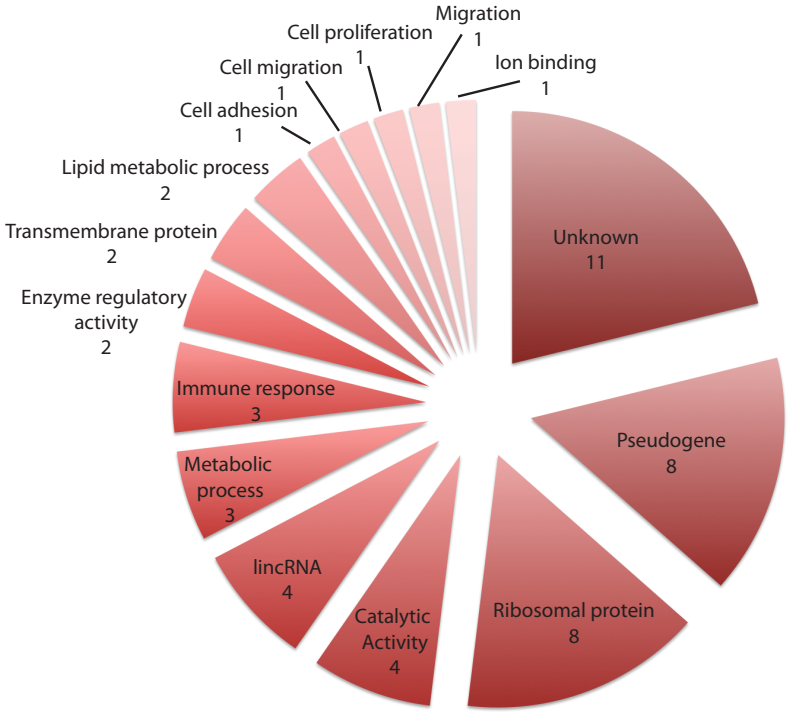



**Figure 2. 4 Differential expression of genes with p < 0.01 when comparing *Matt*$^{ma/ma}$ to C57BL/6J mice.** Transcripts from whole transcriptome sequencing were mapped back onto the reference genome using the Integrated Genome Browser and described according to their Gene Ontology-associated biological processes.

### 2.3.3. Analysis and identification of murine *Tmem79* gene

Further investigation of *Tmem79* was carried out by performing conventional Sanger sequencing of PCR products. The presence of the c.840C>G mutation at chromosomal position chr3: 88,136,485 was confirmed in homozygous $Matt^{ma/ma}$ mice, within exon 3 of *Tmem79*. This mutation was not present in wildtype C57BL/6J or $Flg^{ft/ft}$ mice (Figure 2. 6a). Genomic DNA from the original matted mice from Jackson Laboratory and $Matt^{ma/ma}Flg^{ft/ft}$ double mutants (DM) maintained by our collaborators from Trinity College both harbored the same mutation. It was decided that the mutant allele should be called $Matt^{ma}$.

*Tmem79* is located on mouse chromosome 3qF1 and spans 5,781 bp of genomic DNA. With a total of four exons, the first exon is untranslated and the other three encode for a 391 amino acid protein with a predicted molecular weight 43.5 kDa. Protein analysis predicts 5 transmembrane domains and the mutation p.Y280* truncates TMEM79 before the third transmembrane domain (Figure 2. 5).

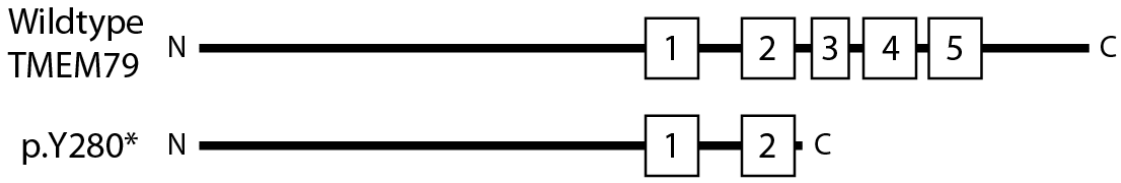



**Figure 2. 5 Bioinformatics protein analysis prediction of TMEM79.** *Tmem79* is predicted to encode for a 5 transmembrane protein and the *matted* mutation p.Y280* truncates TMEM79 between the second and third transmembrane domain (numbered rectangles = transmembrane domains).

### 2.3.4. Mouse genotyping assay

A restriction enzyme digest was designed to enable quick genotyping of mice. A single nucleotide base change from C to G occurs within the exon 3 of *Tmem79* gene. It also conveniently occurs within a restriction site for *CviQI*, which recognizes a G$^{v}$TAC sequence. There are two *CviQI* restriction digest sites within

wildtype (C57BL/6J) genomic exon 3 sequence, resulting in three fragments: 380bp, 186bp and 78bp. Within the homozygous *Matt*$^{ma/ma}$ mice, the C to G change results in one less restriction site, giving only two fragments: 458bp and 186bp (Figure 2. 6b).

### 2.3.5. Protein expression

Using a commercially available rabbit polyclonal antibody against the C-terminus of human TMEM79, our collaborators in Dublin performed an immunoblot on protein extracts from mouse epidermis, and a single band of ~49kDa was seen in C57BL/6J and *Flg*$^{ft/ft}$ mice which was absent from *Matt*$^{ma/ma}$ and *Matt*$^{ma/ma}$*Flg*$^{ft/ft}$ double mutants (DM) (Figure 2. 6c), confirming that the p.Y280* mutations results in undetectable levels of the full-length protein.

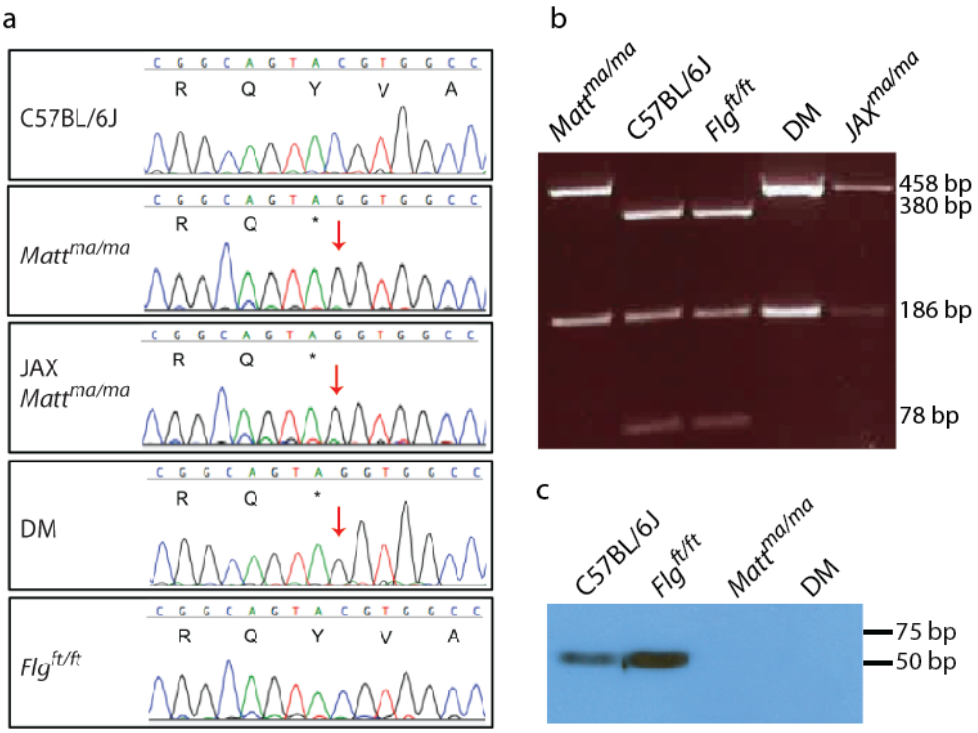



**Figure 2. 6 Further confirmation of the *matted* mutation in various mouse strains.** Chromatograms of sequenced PCR products showing c.840C>G in the isolated *Matt*$^{ma/ma}$, original DM from Jackson Laboratories (JAX) and the DM from Ireland. This mutation was not seen in C57BL/6J or *Flg*$^{ft/ft}$ (a). Genotyping assay utilizing the single base change which abolishes one of the *CviQI* restriction sites within the amplification product (b). Immunoblotting protein extracts from mice skin shows complete absence of protein within *Matt*$^{ma/ma}$ homozygous murine neonatal epidermis (c).

### 2.3.6. Expression analysis of *Tmem79*

Murine *Tmem79* cDNA was amplified using primers that spanned exon 3 and 4. A faint band was detected in *Matt*$^{ma/ma}$ samples compared to that seen in C57BL/6J (Figure 2. 7a). In order to quantify *Tmem79* transcript levels in the epidermis of either mouse strain, a TaqMan® assay that spanned exons 3 – 4 was performed. *Matt*$^{ma/ma}$ mice had 7584 copies of total *Tmem79* while in comparison C57BL/6J mice had 36840 copies. Errors bars in the graph show the standard deviation across three replicates for *Matt*$^{ma/ma}$ (374.78) and C57BL/6J mice (4806.5; Figure 2. 7b). This shows a nearly 5–fold decrease in *Tmem79* message in *Matt*$^{ma/ma}$, suggesting that the p.Y280* nonsense mutation results in post-transcriptional degradation of the mutant *ma* transcript.

The same TaqMan® assay was then performed on a commercially available multiple murine tissue array, and, although high amounts of *Tmem79* message was seen in the stomach and bladder, surprisingly low amounts were seen in the skin sample (Figure 2.8), which was a contradiction to the high amounts seen in control C57BL/6J epidermis (Figure 2. 7). According to the company that supplied the panel, the cDNA panel was supposed to be normalized to *Gapdh* so that all samples were the same, but a standard *Gapdh* assay showed that the samples were not equal across all tissue types (Figure 2.9), which might contribute to the false negative results for *Tmem79* in murine skin.

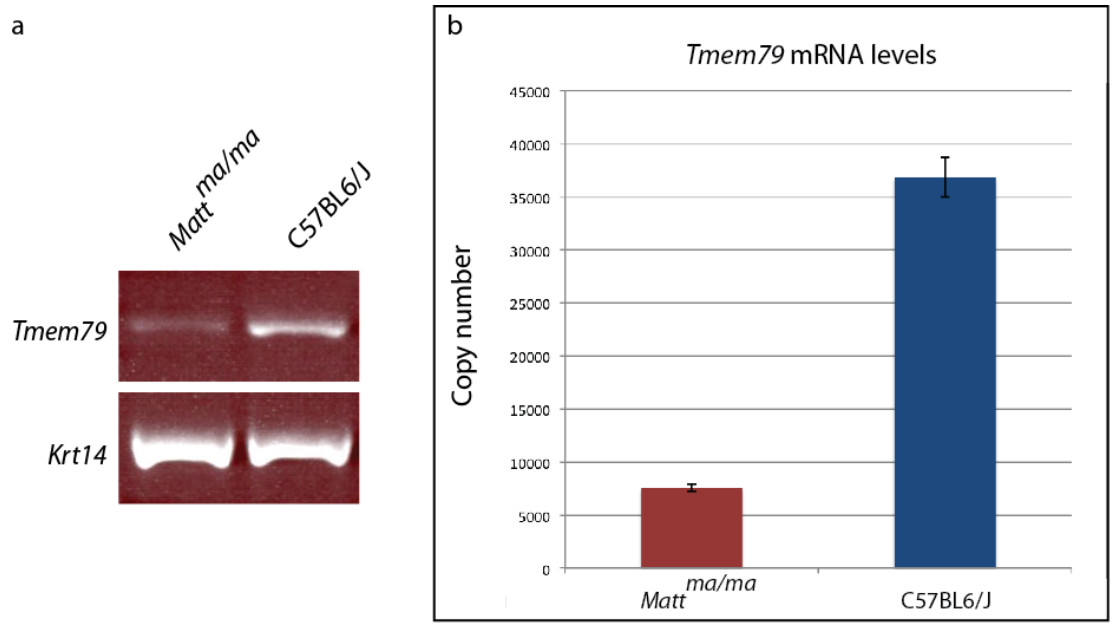



**Figure 2. 7 Investigation into *Tmem79* transcript levels in *Matt*$^{ma/ma}$ mice.** RT-PCR analysis of mouse skin cDNA showed a decrease in *Tmem79* transcript levels, using primers spanning exon 3 – 4 and *Krt14* was used as a positive control (a). Quantification of *Tmem79* using a Taqman® Gene Expression Assay, spanning exon 3 – 4 showing a decrease of nearly 5-fold for *Tmem79* message in *Matt*$^{ma/ma}$ mice compared to C57BL6/J. Error bars show standard deviation (three animals per genotype) (b).

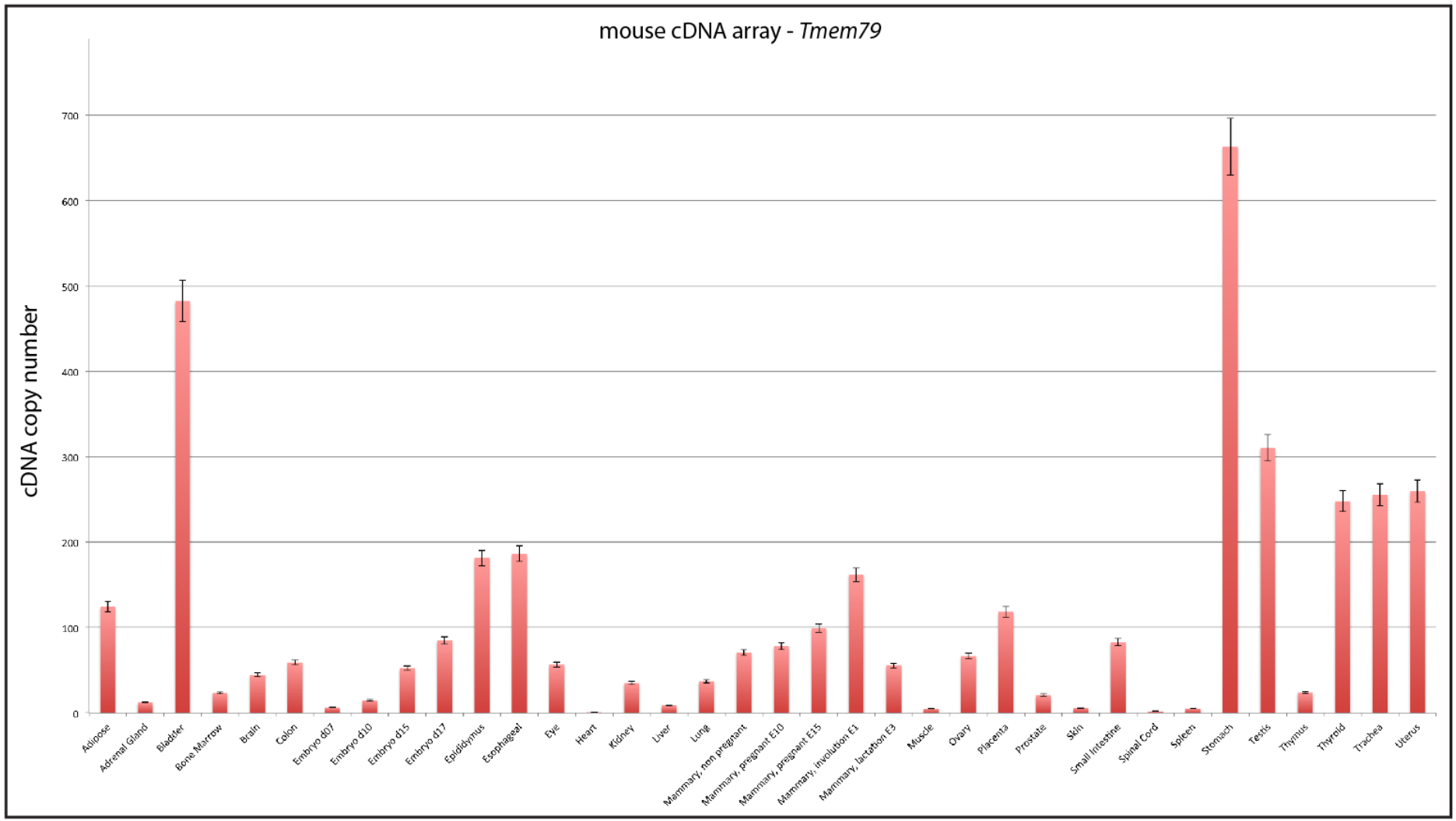



**Figure 2. 8 Multiple tissue expression analysis of *Tmem79*.** Multiple tissue cDNA arrays for murine tissues showing unexpectedly low levels of *Tmem79* in the skin across three replicates. Error bars show standard deviation across 3 multiple tissue arrays.

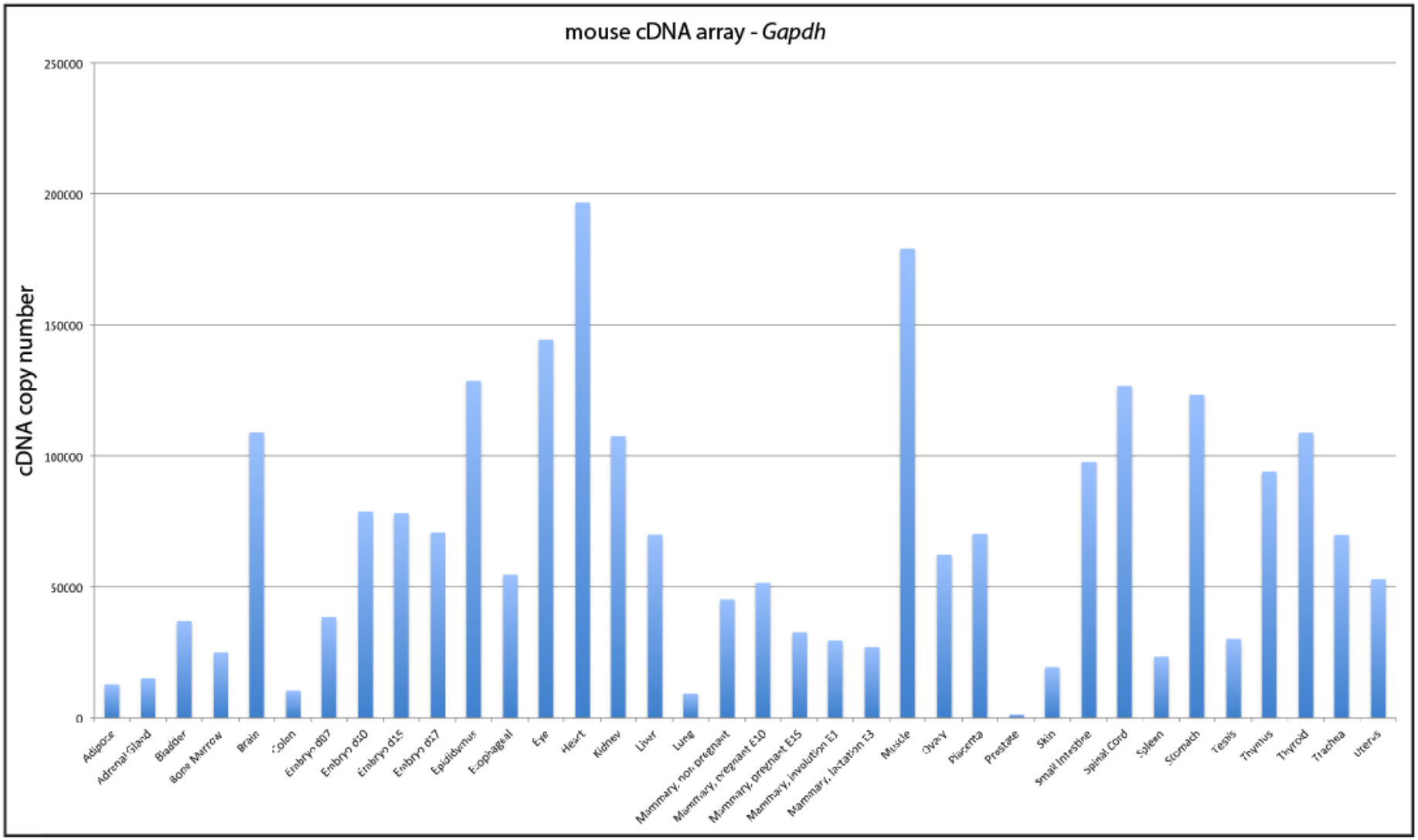



**Figure 2. 9 Multiple tissue expression analysis of *Gapdh*.** Inconsistent amounts of *Gapdh* seen within multiple tissue cDNA array.

### 2.3.7. Analysis of TMEM79 localization and lipid patterns in *Matt*$^{ma/ma}$ mice

Immunofluorescence microscopy of dorsal epidermis from *Matt*$^{ma/ma}$ and C57BL/6J showed that TMEM79 was localized to the upper granular layers in the C57BL/6J mice, but was absent in the *Matt*$^{ma/ma}$ animals (Figure 2. 10a, b; green). Expression was also seen in the hair follicle of C57BL/6J mice (Figure 2. 10a; white arrow) compared to its absence in *Matt*$^{ma/ma}$ (Figure 2. 10b).

Staining with lipophilic Nile Red dye revealed highly ordered stacked of cornified cell envelopes in the stratum corneum of C57BL/6J animals, but these seemed to be disorganized in *Matt*$^{ma/ma}$ mice (Figure 2. 10c, d). There appeared to be a buildup of lipid at the SG/SC interface, with less Nile Red staining in the SC of *Matt*$^{ma/ma}$ mice compared to C57BL/6J.

### 2.3.8. Lipid analysis of *Matt*$^{ma/ma}$ and C57BL/6J skin

Thin layer chromatography (TLC) was performed on skin extracts from *Matt*$^{ma/ma}$ and C57BL/6J dorsal epidermis to identify any differences in lipid content as suggested from Nile red staining. 1D and 2D TLC separation did not reveal any differences in the migration or abundance of phospholipids, fatty acids, polar or non-polar lipids between *Matt*$^{ma/ma}$ and C57BL/6J (Appendix VI). There appears to be a small difference in the amount of non-hydroxy ceramides, but it could be due to an artifact bias, as the overabundance of total ceramides seems to have affected separation of the different ceramide classes.

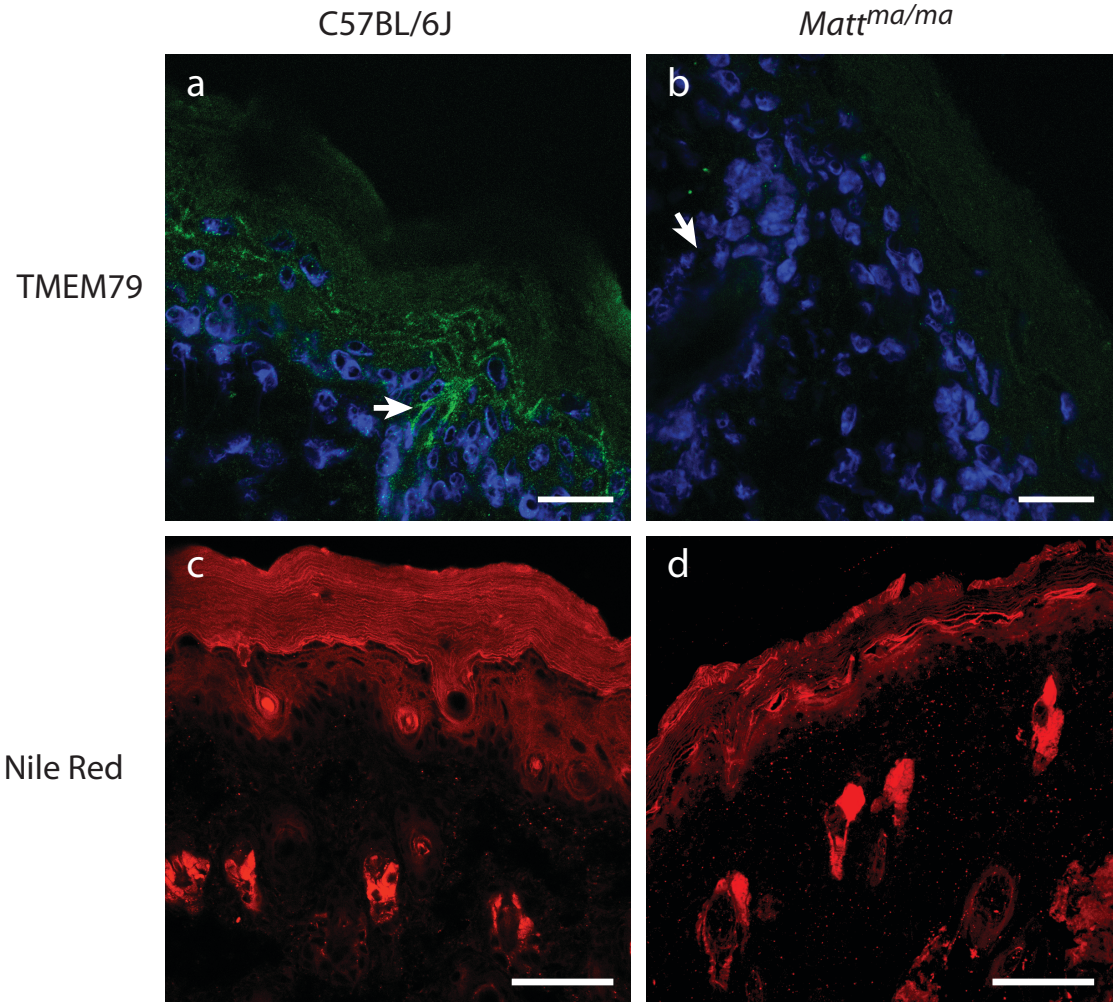



**Figure 2. 10 Analysis of TMEM79 localization and lipid patterns in *Matt*$^{ma/ma}$ dorsal epidermis.** Confocal microscopy of mouse dorsal epidermis stained with rabbit anti-TMEM79 polyclonal antibody (green) showing absence of TMEM79 in *Matt*$^{ma/ma}$ mice (b), compared to expression in C57BL/6J wildtype mice (a). Strong expression of TMEM79 can be seen in murine C57BL/6J hair follicles (white arrow). Nuclei were stained with DAPI (blue; a, b). Nile red staining of wildtype mice show highly ordered stacks of cornified cell envelopes (c) but were highly disorganized in *Matt*$^{ma/ma}$ (d). Scale bar = 50 µm.

## 2.4. DISCUSSION

The recent discovery of a frameshift mutation in the FLG gene in flaky tail mice has led to the use of *Matt*$^{ma/ma}$*Flg*$^{ft/ft}$ double mutant (DM) as a model for skin barrier disease. Several publications have used the DM as an animal model for skin barrier abnormality studies (Moniaga et al., 2010; Moniaga et al., 2013; Oyoshi et al., 2009; Scharschmidt et al., 2009). The stratum corneum of the DM has been shown to be exceptionally permeable to externally applied molecules, showing that the animals have a severe barrier abnormality (Scharschmidt et al., 2009) as well as abnormal lamellar granule (LG) secretion, resulting in a buildup of unsecreted LG at the SG-SC interface. There appears to be normal LG formation, but the process of LG secretion is hampered, leading to a decrease in secreted material and lamellar bilayers in these DM mice. Barrier abnormality in DM mice results in an overall elevated immune response driven by both $T_H2$-mediated inflammation (Moniaga et al., 2013; Oyoshi et al., 2009) and $T_H1/T_C1$-mediated responses (Moniaga et al., 2010). Since these experiments were, however, performed on DM animals, it cannot be determined whether the resulting inflammatory profile is the result of the *ft* or *ma* mutation. The only publication to date, working on *Flg*$^{ft/ft}$ only mice, shows abnormal keratin aggregation within the lower SC layers, suggesting that these mice are more susceptible to mechanical stress (Kawasaki et al., 2012). In addition, *Flg*$^{ft/ft}$ mice are susceptible to percutaneous sensitization and exhibit an exaggerated dermatitis-type response. Interestingly, the paper went on to comment that the *Flg*$^{ft/ft}$ animals showed normal LG secretion and lamellar bilayers at the SG-SC interfaces. From this, we can deduce that while *Flg* does play an important role in skin barrier formation, the cause of abnormal LG secretion could be attributed to the *ma* gene.

The breeding strategy we used allowed isolation of the *Matt*$^{ma/ma}$ mutation, and mapping primers led us to identify a critical homozygous region of 1.05 Mb on chromosome 3. Traditionally, disease genes have been identified using karyotyping, linkage analysis, homozygosity mapping, copy number variation analysis and SNP-based association analysis. Homozygosity mapping was chosen initially, because prior genetic linkage studies showed that the *matted* gene was approximately 1.5 cM away from *Flg,* and would be relatively easy to find. However, that was not the case and next generation sequencing had to be employed. If next generation sequencing has been carried out initially, it would have reduced the disease gene identification from two-step, to a one-step approach (Gilissen et al., 2012).

A shortlist of candidate genes was identified from whole transcriptome sequencing data, and *Transmembrane79 (Tmem79)* was shown to have a nonsense mutation that would generate a non-functional protein. Differential gene expression data identified a significant decrease in the amount of mRNA transcripts in the *Matt*$^{ma/ma}$ mouse, suggesting that the mutant *Tmem79* transcript was undergoing nonsense-mediated decay. Expression analysis of *Matt*$^{ma/ma}$ epidermis showed a similar decrease in *Tmem79* mRNA compared to C57BL6/J animals. The results from differential gene analysis across the entire genome of the *Matt*$^{ma/ma}$ mouse only identified *Tmem79* within the epidermal differentiation complex on chromosome 3 as showing any difference in expression. Taking this into consideration, it would perhaps have minimized the amount of work carried out for homozygosity mapping if whole transcriptome analysis were carried out first. Homozygosity mapping did highlight the critical region of genes and allowed analysis of fewer genes within the entire EDC, but a combination of data from differential gene analysis and SNPs/InDels identified by whole transcriptome

analysis would have flagged up *Tmem79* initially, which could then be confirmed with Sanger sequencing.

The data generated from whole transcriptome analysis of *Matt*$^{ma/ma}$ reflected a huge increase in the expression of genes involved in the immune response. The majority of these immune-related genes code for histocompatibility class 2 molecules, which are involved in antigen presentation by phagocytes. Other transcripts coding for proteins involved in inflammatory pathways that were upregulated in the *Matt*$^{ma/ma}$ animal are caspase 1 and lysozyme 2. This correlates well with the atopic phenotype of the *Matt*$^{ma/ma}$ mouse, showing that the apparent skin barrier defect results in a hypersensitized immune system. The Wdfy1 (WD repeat and FYVE domain containing 1) message showed a four-fold increase in expression in the *Matt*$^{ma/ma}$ mouse, which may be significant since the Fyve domain is able to target proteins to membrane lipids via an interaction with phosphatidylinositol-3-phosphate, a possible interaction partner with the TMEM79. CD1d1 and lipoprotein lipase showed a close to 2-fold increase in expression in *Matt*$^{ma/ma}$ mice, with the former protein belonging to a family of glycolipid antigen-presenting MHC-like molecules with a role in Th1 and Th2 type responses. The latter protein is involved in the enzymatic hydrolysis of triglycerides. Two genes encoding for lipid-related proteins showed decreased expression in the *Matt*$^{ma/ma}$ mouse, and they are fatty acid desaturase 6 (Fads6) and epidermal fatty acid binding protein 5 (Fabp5). The overall landscape of gene expression in *Matt*$^{ma/ma}$ suggests that TMEM79 is required for normal lipid homeostasis within the epidermis, and defects in lipid homeostasis alter the functionality of skin barrier, leading to an elevated immune response. In our experiments, we used epidermis from 4-day old animals in order to evaluate the normal-state of murine epidermis before the overt reactive inflammatory

phenotype of the *Matt*$^{ma/ma}$ adult mice developed. It is possible that the inflammatory pathways/immune responses of the *Matt*$^{ma/ma}$ animal could change over time, and it would be interesting to look at the immune profile of the *Matt*$^{ma/ma}$ mouse at different ages. An attempt to perform pathway analysis using STRING (v9.05) was performed using the list of genes with differential expression, but no obvious patterns or interactions between the genes were discerned. This could be due to the complexity of the challenges in trying to annotate these genes as most of them were "unknown", and the difficulties involved in predicting interactions between such a wide range of genes which may be under-represented in various pathways.

Lipid analysis did not show any differences in epidermal lipid contents (phospholipids, fatty acids, polar or non-polar lipids and ceramides) between the *Matt*$^{ma/ma}$ and C57BL/6J animals, even though nile red staining showed disordered cornified cell envelope stacks in the stratum corneum. *Flg*$^{ft/ft}$ mice have been shown to have an abnormal increase in levels of ceramide and cholesterol of the SC, which affects the hydrophobicity of the SC and the penetration of haptens/allergens (Kawasaki et al., 2012). As shown in a previous paper (Scharschmidt et al., 2009), there could be a build up of entombed lamellar bodies at the SG-SC interface in *Matt*$^{ma/ma}$ mice, and whilst the overall lipid profile may appear normal, they may not be at the appropriate sub-cellular locations for lamellae formation.

The *Matt*$^{ma/ma}$ mouse has a mutant TMEM79 protein missing two thirds of its C-terminus. The location of *Tmem79* within the EDC suggests that it is involved in the maturation or differentiation processes of the epidermis, and could be engaged with the biosynthesis and/or transport of epidermal lipids. This process is a vital part of epidermal differentiation between the SG and SC layers, where lamellar

granules fuse with the cell lipid bilayer and contents are extruded to the intercellular spaces to form the lipid component of the cornified cell envelope. The commercial antibody used was designed using a synthetic peptide of the sequence between 364 – 394 amino acids from the C-terminal region of human TMEM79. The mouse TMEM79 protein sequence has 3 three different peptides within this region compared to the human TMEM79 sequence, and this difference might account for the weaker detection of TMEM79 using immunofluorescence of mouse dorsal epidermis.

The matted hair phenotype of the *Matt*$^{ma/ma}$ mouse has been well characterized in an early paper (Jarrett and Spearman, 1957), where the fragility and brittleness of the hair shaft culminates in alopecia and as a result, the hair cycle is significantly shorter than in normal mice. It is possible that the mutant TMEM79 is affecting the release of lamellar granule products such as involucrin, loricin, TGase 1, 2, 3, which have been shown to be expressed at the inner root sheath of the hair canal (Akiyama et al., 2002). This in turn, might well affect the differentiation of the developing hair follicle of the *Matt*$^{ma/ma}$ animal.

By managing to isolate the *Matt*$^{ma/ma}$ strain, we have shown that it presents with a more severe phenotype compared to the *Flg*$^{ft/ft}$ mouse, and this is attributed to the p.Y280* mutation leading to a truncated TMEM79 protein. The resulting hyper-immune response and hair shaft abnormalities seen in the *Matt*$^{ma/ma}$ mouse must be the effect of abberant lipid transport/secretion at the SC-SG junction in the epidermis, or the IRS of the hair follicle.

# CHAPTER 3

# Characterization of human *TMEM79*

## 3.1. AIMS OF CHAPTER

The identification of the novel murine gene *Tmem79* in Chapter 2 of this thesis, has revealed the genetic cause of the mouse mutant *Matt*$^{ma/ma}$. TMEM79 appeared to be affecting epidermal lipids in the *Matt*$^{ma/ma}$ mouse, giving rise to an eczema-like inflammatory skin condition. In this chapter, I attempt to characterize the human ortholog of TMEM79, and identify any involvement with epidermal barrier formation or differentiation.

## 3.2. MATERIALS AND METHODS

### 3.2.1. Identification of human *TMEM79*

Primers were designed to amplify the human *TMEM79* gene on chr1: 156254070-156262234 using similar PCR conditions described in Chapter 2.2.2 (page 48), with annealing temperatures for primer pairs for all exons 54°C, apart from primers used for exon 2.2, which have an annealing temperature of 58°C (Table 3. 1).

| Gene/mRNA | Location on human chr1 | | F primer (5' - 3') | R primer (5' - 3') |
|---|---|---|---|---|
| *TMEM79* | 156254070:156262234 | exon 1 | GCCTGCTGCTACCAAATCT | CTTTACTCCTATGTGATGCTCG |
| | | exon 2.1 | TGGACGGCATTTGATGTC | TGTAGGTCAATAGGCACGAAG |
| | | exon 2.2 | GACGATGCCAACCTGCTG | AGGGTAGGGTCAGAAAAGGC |
| | | exon 3 | TTTCCCTTCACTTGACCTG | GAAATCGGCTTCCATCTG |
| | | exon 4.1 | TACCAGCCATACCCACCAAC | ACGGCTGTCCACCTTTTTC |
| | | exon 4.2 | CAGGGAAAAAGGTGGACA | TAACATCAACCACCGCAC |

**Table 3. 1** PCR and sequencing primers for human *TMEM79* gene.

### 3.2.2. Semi-quantitative RT-PCR of *TMEM79* transcript variants

Exon-specific forward primers were designed for each *TMEM79* transcript variant, each of the forward primers in exon 1 were designed to span the first intron using the same reverse primer located within exon 2. Primers were also designed to identify transcript variant of exon 2, with the expectation of a smaller sized product for transcript variant 2 (425 bp), shown in Table 3. 2**.** A diagrammatic representation of transcript variant primers are shown in Figure 3. 2 (page 90) as blue/green (forward) and red (reverse) arrows. Semi-quantitative RT-PCR was carried out on cDNA obtained from frozen normal human skin (Tayside Tissue Bank, Dundee), human primary keratinocytes (HPK) and HaCaT cells using the same protocol detailed in Chapter 2.2.7 (page 51). Primers amplifying human keratin 14 cDNA were used as a positive control.

| Human cDNA primers | Chromosome position | F primer (5' – 3') | R primer (5' -3') | Expected mRNA size (bp) |
|---|---|---|---|---|
| *TMEM79* exon 1a | chr1:156252766+156255694 | CAAACCCTTGCGAAAAAC | ACATCAAACGGCAGGAAG | 811 |
| *TMEM79* exon 1b | chr1:156254113+156255694 | CTTGCAGAGCCAGCAGGTA | | 847 |
| *TMEM79* exon 2 | chr1:156255013+156255694 | CCAGGATGACAGAACAGGAG | | 683 or 425 |
| Keratin 14 | chr17:39740024-39742699 | CTGGACAAGGTGCGTGCTCTG | CCATTCCTCGGCATCCTTGC | 528 |

**Table 3. 2** Primers used to amplify *TMEM79* transcript variants. Forward primers used were specific for each exon 1 transcript variant. Exon 2 primers are expected to give 425 bp-sized product if exon 2 transcript variant is present.

### 3.2.3. Quantitative PCR of human cDNA tissue array for *TMEM79* transcript variants

Expression levels for human *TMEM79* were quantified across a multiple tissue panel containing cDNA from 48 different human tissues (Origene Technologies, Inc., Maryland, USA) using a TaqMan® Gene Expression Assay (Hs00260483_m1, Life Technologies Corporation, Applied Biosystems, USA). The human multiple tissue cDNA array had been standardized and undergone quality control against human GAPDH with RT-PCR and standard PCR. The *TMEM79* probe spans exon 3 to 4 and gives an amplicon length of 65 bp and will detect all transcript variants. The location of the assay probe within *TMEM79* is Chr1: 156252704-156262234. The DNA standards used in this quantitative assay was a commercially available plasmid containing the full-length wild type *TMEM79* cDNA (Origene Technologies Inc., Maryland, USA). Hs01076300_m1 was used as a gene expression assay for *TMEM79* transcript variant 1a, and a custom assay was designed for *TMEM79* transcript variant 1b (AJWR2C5). Standards for each transcript variant were amplified with the appropriate PCR primers listed in Table 3. 2, cloned into a vector with known size and mass, and the standard curve was calculated using the formulae provided by Applied Biosystems

("http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_042486.pdf").

### 3.2.4. Histological tissue section expression

Human scalp and cheek skin (Tayside Tissue Bank, Dundee) were snap frozen in liquid nitrogen and stored at – 80°C until needed. 8 µm cryosections were cut and processed in an identical manner to that stated in Chapter 2.2.9, page 53. Primary antibodies used and their working dilutions for indirect immunofluorescence on tissue sections are listed in Table 3.3.

Primary antibodies against hair follicle keratins 32, 74 and 75 were a kind gift from Dr Lutz Langbeing (DKFZ, Heidelberg, Germany). For hair follicle staining, frozen scalp sections were stained as above but with an additional permeabilization step using 1% (v/v) NP-40 (AppliChem GmbH, Darmstadt, Germany) prior to blocking with 10% (v/v) goat serum in PBS.

All secondary antibodies were used at 1:1000 dilution in 10% (v/v) goat serum in PBS. They include AlexaFluor® Goat-anti-mouse 488, AlexaFluor® Goat-anti-rabbit 594, AlexaFluor® Goat-anti-rabbit 488, AlexaFluor® Donkey-anti-rabbit 488 and AlexaFluor® Goat-anti-mouse 594 and AlexaFluor® Goat-anti-guinea pig 594 (Invitrogen, Life Technologies Corporation, Maryland, USA).

| | Host Species | Dilution for IF | Clone | Catalogue no. | Source |
|---|---|---|---|---|---|
| β-actin | Mouse | 1:500 | AC-74 | A2228-200UL | Sigma-Aldrich, Missouri, USA |
| β-COPI | Rabbit | 1:200 | -- | ab2899 | Abcam plc, Cambridge, UK |
| COPII | Rabbit | 1:200 | -- | PA1-069A | Thermo Scientific, Illinois, USA |
| Corneodesmosin | Mouse | 1:500 | F28-27 | A kind gift from Guy Serre, University of Toulouse, France | |
| Desmoglein 1 | Mouse | 1:100 | 27B2 | MCA2271 | AbD Serotec, Kidlington, UK |
| Desmoplakin 1 & 2 | Mouse | 1:100 | DP2.15 | CBL173 | Millipore, California, USA |
| EEA1 | Rabbit | 1:500 | -- | ab2900 | Abcam plc, Cambridge, UK |
| Filaggrin | Mouse | 1:100 | 15C10 | 1mL NCL-FILAGGRIN | Novocastra Laboratories Ltd, Newcastle, UK |
| Involucrin | Mouse | 1:100 | SY5 | ab68 | Abcam plc, Cambridge, UK |
| Kallikrein 8 | Mouse | 1:30 | 2C6 | TA506171 | Origene Technologies, Inc., Rockville, Maryland, USA |
| Keratin 5 | Mouse | 1:100 | RCK102 | ab9005 | Abcam plc, Cambridge, UK |
| Keratin 10 | Mouse | 1:20 | DE-K10 | M700201-2 | Dako UK Ltd, Cambridgeshire, UK |
| Keratin 14 | Mouse | 1:5000 | LL001 | SC53253 | Santa Cruz Biotechnology Inc., Heidelberg, Germany |
| Keratin 17 | Mouse | 1:300 | CK-E3 | C9179 | Sigma-Aldrich, Missouri, USA |
| Keratin 32 | Guinea pig | 1:1000 | Ha2.1 | A kind gift from Dr Lutz Langbein, DKFZ, Heidelberg, Germany | |
| Keratin 74 | | 1:2000 | CK T4.1 | | |
| Keratin 75 | | 1:2000 | Bax-1 | | |
| LAMP2 | Mouse | 1:100 | H4B4 | ab25631 | Abcam plc, Cambridge, UK |
| LC3 | Mouse | 1:500 | 4E12 | M152-3 | MBL International, Japan |
| Loricrin | Goat | 1:100 | C-13 | sc-51130 | Santa Cruz Biotechnology Inc., Heidelberg, Germany |
| Myc-tag | Mouse | 1:500 | 9B11 | 2276 | Cell Signaling Technology, Massachusetts, USA |
| Rab5 | Rabbit | 1:500 | -- | ab18211 | Abcam plc, Cambridge, UK |
| TGN46 | Sheep | 1:100 | -- | AHP500GT | AbD Serotec, Kidlington, UK |
| TMEM79 | Rabbit | 1:100 | RB26498 | AP5286b | Abgent, California, USA |

**Table 3. 3** Table of primary antibodies used for indirect immunofluorescence.

### 3.2.5. Immuno-gold labeling of ultra-thin skin sections

The Microscopy Unit, University of Dundee (John James), performed the preparation and ultra-thin cryosectioning of normal human abdominal skin (Tayside Tissue Bank, Dundee). The methods employed are described below.

Human skin sections were trimmed down to ~1 $mm^2$ cubes and fixed in 4% (w/v) paraformaldehyde with 0.2 M piperazine-N, N'-bis (2-ethanesulfonic acid) (PIPES buffer) at 37°C and transferred into 2.1 M sucrose for storage at 4°C. Dermis was trimmed using a Leica EM FCS ultramicrotome to remove as much as possible from the epidermis. Each 90 nm ultra-thin section was retrieved from the diamond knife onto a hexagonal 100-mesh copper grid coated with pioloform and carbon. Grids were suspended within a 1:1 solution of 2.3 M sucrose and 2% (v/v) methylcellulose droplet on a stainless steel loop and allowed to dry before suspending on water for short-term storage.

Each copper grid was washed twice in PBS at 1 min each, and blocked in 0.5 % (w/v) fish skin gelatin in PBS (FSG/PBS) for 10 mins. The grids were then transferred into primary antibody in FSG/PBS for 30 mins, before washing in PBS three times, 5 mins each. The appropriate gold conjugate secondary antibody was applied for 20 mins, and grids were washed six times in PBS for 5 mins each, then a further 10 ten washes in distilled water for 1 min each, before counterstaining the sections in 2% (v/v) methyl cellulose and 3 % (w/v) uranyl acetate. Sections were examined in a Jeol 1200 Ex electron microscope, images photographed on digital imaging plates and scanned in a high resolution imaging plate scanner (DITABIS, Pforzheim, Germany).

Antibodies used and their respective secondary antibodies for immuno-gold labelling were diluted in FSG/PBS:

1:4 dilution of anti-TMEM79 with 1:50 dilution of goat anti-rabbit 10nm Au.

1:4 dilution of anti-CDSN with 1:50 dilution of goat anti-mouse 15 nm Au.

1:2 dilution of anti-KLK8 with 1:50 dilution of goat anti-mouse 15 nm Au.

All gold conjugate secondary antibodies used were from British Biocell International, Cardiff, United Kingdom.

### 3.2.6. Protein extraction from skin samples

Normal frozen human samples (Tayside Tissue Bank) were cut into small pieces and ground into powder using a mortar and pestle. The powder was then homogenized on ice in a Dounce homogenizer with lysis buffer containing; 7 M Urea, 0.1 M dithiothreitol, 0.05% (v/v) Triton X, 25 mM NaCl and 20 mM HEPES, pH 8.0. After sufficient homogenization, the extract was transferred to an eppendorf tube and incubated at 37°C for an hour.

### 3.2.7. Cloning of TMEM79-GFP constructs

Using a commercially available myc-tagged *Homo sapiens* TMEM79 (RC202998, Origene Technologies, Inc., Rockville, Maryland, USA), as a template, full-length transcripts were amplified using primers incorporating restriction enzymes sites to facilitate downstream cloning. The sequence used as a template was transcript variant 1a. Primers were designed and conditions were optimized (Table 3. 4). Following amplification, 1.2 µL of PCR product was cloned into pCR®2.1-TOPO vector using a TOPO TA Cloning kit (Life Technologies Corporation, Invitrogen, Maryland, USA) according to manufacturer's instructions. Ligation products were transformed into DH5α cells, and spread on a LB ampicillin plate containing 40 µg/mL X-gal and 0.2 mM IPTG and incubated at 37°C overnight. Selected white colonies were suspended in 50 µL of LB broth and colony-PCR was carried out using the initial amplification primers. After confirming the correct band size on an

agarose gel, the colony was grown in 2 mL liquid culture overnight at 37°C with the appropriate antibiotics. Plasmids were isolated from the culture using a Qiaprep Spin Miniprep kit (Qiagen, Gmbh, Hilden, Germany) according to manufacturer's instructions and plasmid DNA sent for sequencing. Sequences for each plasmid were checked for base changes and insert orientation. For subcloning into pEGFP-C1 and pEGFP-N1 vectors, cloned products were digested with 1 unit of the appropriate restriction enzyme (New England Biolabs, Maryland, USA), in a total reaction volume of 20 µL. Controls for vector only, single digest and double digest were carried out. The pEGFP-C1 and pEGFP-N1 vectors (Clontech Laboratories Inc., Takara Bio Company, California, USA) were similarly digested and the vectors were treated with 1 unit of Calf Intestinal Alkaline Phosphatase (CIP) (New England Biolabs, Maryland, USA) at 37°C for 3 hours. Digested products were run on an agarose gel, the band of interest extracted and purified using QIAquick PCR-Purification kit (Qiagen, Gmbh, Hilden, Germany). Ligation with T4 DNA ligase (New England Biolabs, Maryland, USA) was carried out at 16°C for 12 hours. DH5α cells were transformed with the ligated product and plated on LB kanamycin plates overnight at 37°C. Several colonies per clone were selected and 2 mL starter-cultures were grown overnight at 37°C. The plasmids were isolated with Qiaprep Spin Miniprep kit. Each plasmid was double-digested with their respective restriction enzymes to ensure the presence of the correct sized insert before transformation and a final plating on kanamycin-infused LB plates overnight at 37°C. Two colonies were selected for each fusion protein and inoculated into 100 mL of LB media with the appropriate antibiotics (in a flask that was at least five times the volume of liquid) and left overnight to grow at 37°C to ensure maximal bacterial growth. Plasmids were harvested the next day with QIAprep Plasmid Maxi Kit (Qiagen, Gmbh, Hilden, Germany). Small stocks of 1 - 2

mL of culture were mixed with 80 % (v/v) glycerol and stored at -80°C for future use.

| Primer name | Primer sequence (5' - 3') | RE site | Size (bp) |
|---|---|---|---|
| TMEM79-FL-GFPN-F | ATCCGGTACCGAGGAGATCT | *BglII* | ~1220 |
| TMEM79-FL-GFPN-R | TGCTCGACCGGTCGCGTACGC | *AgeI* | |
| TMEM79-FL-GFPC-F | ATCCGGTACCGAGGAGATCT | *BglII* | |
| TMEM79-FL-GFPC-R | GCT CCT GCA GCC GCG TAC GCT AGC C | *PstI* | |

**Table 3. 4** Primers used in amplifying protein constructs for full-length (FL) TMEM79, with introduced restriction enzyme (RE) sites and their expected sizes.

### 3.2.8. Site-directed mutagenesis

The full-length TMEM79/pEGFPC1 clone was used as a template for site-directed mutagenesis to introduce the mutation c.T847A (p.Y283*), the human equivalent of the matted mouse mutation p.Y280*. Primers were designed taking several factors into consideration; length should be between 25 – 45 bp with a melting temperature of >78°C, the desired mutation to be in the middle of the primer, a minimum GC content of 40 % and HPLC-purified (Table 3. 5). The mutant strand was synthesized in a 50 µL amplification reaction with 2.5 U/µL PfuUltra Fusion HotStart DNA Polymerase and 1x PfuUltra II reaction buffer (Agilent Technologies, California, USA), 125 ng each of forward and reverse primer and 2.5 mM of each dNTP (Bioline). Conditions for amplification used were: one cycle of 95°C and 16 cycles of 95°C for 30 seconds, 55°C for 1 minute and 68°C for 6 minutes. The reaction was allowed to cool to 4°C before the addition of 1 µL of restriction enzyme, *DpnI*, then incubated at 37°C for 1 hour to ensure complete digestion of methylated parental strands. DH5α cells were transformed with the digested product and plated on kanamycin-infused LB plates overnight. Colony PCR was carried out the next day on individual colonies to ensure that the correct sized insert was present and overnight cultures started up in 2 mL of LB with

kanamycin. Colonies were isolated, purified and sequenced to check for the presence of the mutated base.

| Primer name | Primer sequence (5' - 3') | RE site | Size (bp) |
|---|---|---|---|
| t3904a_F | GGAGATCCACCGGCGATAAGTGGCCCAGT | BglII | ~1220 |
| t3904a_R | ACTGGGCCACTTATCGCCGGTGGATCTCC | PstI | |

**Table 3. 5** Primers designed for site-directed mutagenesis of the *Matt*$^{ma/ma}$ mutation, c.T847A (p.Y283*; base change highlighted in red) with restriction enzyme (RE) sites and expected size.

### 3.2.9. Cell culture

HaCaT keratinocytes (Boukamp et al., 1988) were grown and maintained in antibiotic-free Dulbecco's Modified Eagle Medium (DMEM) (Gibco, Life Technologies Corporation, Maryland, USA) containing 1.8 mM calcium chloride, 4.5 g/L glucose, L-glutamine and pyruvate, together with 10% Fetal Calf Serum (Hyclone, Thermo Scientific Inc., Illinois, USA). Normal human keratinocytes were obtained from CELLnTECH and cultured in defined serum-free growth medium CnT57 (TCS Cellworks Ltd, Buckingham, UK). All cell lines were maintained in a monolayer in a 5% $CO_2$ environment at 37°C, with media being changed every 2 days until approximately 80 - 90% confluency.

### 3.2.10. Transient overexpression of TMEM79 constructs

Cells were grown in 6-well plates (Greiner Bio-one, MEDTEC Europe, Stuttgart, Germany) containing 4 sterile 13 mm glass coverslips (VWR International Inc., Chicago, USA) per well. Transfection of TMEM79-fusion protein constructs was performed using Lipofectamine™ 2000 (Invitrogen, Life Technologies, Maryland, USA). The volumes of Lipofectamine and Optimem used were optimized to 6 µL and 100 µL per well respectively. The mixture of Lipofectamine + Optimem and Lipofectamine + Plasmid DNA (0.2 µg – 2.0 µg) was incubated for 20 minutes at

room temperature. Whilst waiting, media in each well was changed and after incubation, the complex added to each well and mixed by swirling gently. Cells were incubated at 37°C for the appropriate incubation time (24 – 72 hours). All plasmids used for transfection are shown in Figure 3. 1.

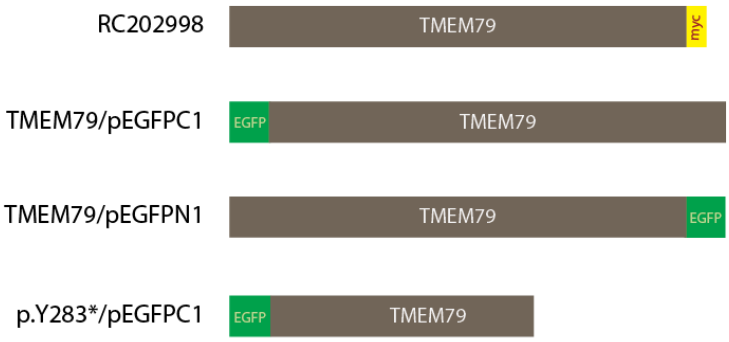



**Figure 3. 1 Overexpression assays using various TMEM79 constructs.** TMEM79-fusion constructs used for transient overexpression with different tags, myc (yellow) and EGFP (green). The sequence from RC202998 was used as the template for the generation of the other constructs.

### 3.2.11. Protein extraction from transfected monolayers

Transfected HaCaT cells grown in 6-well plates were washed twice with PBS on ice. Cells in each well were lysed with 100 µL of lysis buffer containing 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 % (v/v) Triton X-100, 1% (w/v) sodium deoxycholate, 0.1 % (w/v) SDS, 1 mM EDTA, 1x protease inhibitors (Sigma-Aldrich, Missouri, USA) and 2 µL benzonase (Merck KGaA, Darmstadt, Germany) before collecting and shaking horizontally for 10 minutes at room temperature. Each lysate was centrifuged at 12,000 rpm at 4°C for 8 minutes and the supernatant collected and stored at -80°C until needed.

Samples were denatured at 70°C for 10 minutes with NuPage® LDS sample buffer containing NuPage® sample reducing agent and 10 µg/µL of total protein per sample were run on a Novex® NuPage® 4 – 12 % Bis-Tris gel at 200 V for 1 hour using NuPage® MOPS SDS running buffer. Separated proteins were transferred onto nitrocellulose paper, using an XCell SureLock® Mini-Cell (all components

stated above are from Life Technologies Corporation, Applied Biosystems, USA) at 30 V for 50 minutes. Nitrocellulose paper was washed in deionized water and stained with PonceauS (Sigma Aldrich, Life Technologies Corporation, Maryland, USA) for 5 minutes and washed in water to visualize bands. Molecular weight standards (SeeBlue® Plus2 Prestained; Life Technologies Corporation, Applied Biosystems, USA) were marked with a pencil, and nitrocellulose was blocked with 5 % (w/v) Marvel (Chivers Ireland Ltd, Coolock, Dublin) in 0.2 % (v/v) Tween-20 in Tris-Buffered Saline (TTBS) overnight at 4°C before incubation with the primary antibody diluted to the appropriate concentration in 0.06% (v/v) TTBS for 1 hour at 37°C. The blot was washed once with Tris Buffered Saline (TBS), thrice with TTBS and a final wash with TBS, for 5 minutes each wash at room temperature. Incubation with the appropriate secondary antibody (LI-COR Biosciences Ltd, Cambridge, United Kingdom) was carried out for 1 hour before washing (same as above), and imaging carried out with an Odyssey imaging system (LI-COR Biosciences Ltd, Cambridge, United Kingdom).

Primary antibodies and dilutions used for Western blotting:

Anti-EGFP mouse monoclonal at 1:1000 (Roche Diagnostic Ltd, West Sussex, United Kingdom)

Anti-TMEM79 rabbit polyclonal at 1:1000 (Abgent, California, USA)

All LI-COR secondary antibodies were used at a dilution of 1:5000.

### 3.2.12. Indirect immunofluorescence of monolayer cultures

HaCaT cells grown in monolayers on coverslips were fixed at room temperature in 1:1 (v/v) methanol:acetone (stored -20°C) and stained using the same method described earlier (Chapter 2.2.9, page 53). All antibodies used for labeling are listed in Table 3.3 (page 79).

### 3.2.13. Transmission electron microscopy

HaCaT cells were grown on gridded coverslips (Cellocate, discontinued, Eppendorf, Hamburg, Germany) in DMEM with 10% (v/v) FCS at 37°C in sterile conditions. Transfection using TMEM79/pEGFPC1 was carried out according to the protocol detailed in Chapter 3.2.9. Transfected cells were identified and highlighted for easy identification on the labeled grid by fluorescence microscopy. The coverslips were fixed in 4% paraformaldehyde, 2.5% glutaraldehyde in 0.1 M sodium cocodylate buffer at pH 7.3 and post-fixed in 1% aqueous osmium tetroxide. The following steps were performed by John James (Microscopy Unit, University of Dundee). Samples were dehydrated in graded ethanols and propylene oxide and embedded in Durafan resin (Sigma-Aldrich, Missouri, USA). Sections were cut on a Leica Ultracut UCT ultramicrotome, and collected on hexagonal 200mesh copper grids then stained with uranyl acetate and lead citrate. Sections were examined in a Jeol 1200 Ex electron microscope and photographed on digital imaging plates which were then scanned in a high resolution imaging plate scanner (DITABIS, Pforzheim, Germany).

## 3.3. RESULTS

### 3.3.1. Identification and quantification of human *TMEM79* and transcript variants

The syntenic human ortholog *TMEM79* is located on chromosome 1q23.1, with an identical exon-intron organization to the murine *Tmem79* gene. According to the UCSC Genome Browser (Meyer et al., 2013), there are three possible transcript variants of *TMEM79*: two transcript variants of exon 1; uc010phi.2, uc009wrw3, which will be referred to as transcript variant 1a and 1b respectively, and uc001fod.3, which will be referred to as transcript variant 2 for convenience (Figure 3. 2). Specific PCR primers were designed for *TMEM79* transcript variants to detect their presence in tissues. From cDNA extracted from human epidermis, *TMEM79* transcript variant 1b was found to be the major transcript variant with only a low level of transcript variant 1a and no transcript variant 2. In comparison, human primary keratinocytes in a monolayer culture expressed both transcript variant 1a and 1b in equal amounts. cDNA from HaCaT cells, an immortalized human keratinocyte cell line, produced no detectable band for transcript variant 1a, and very little product for transcript variant 1b (contrast was turned up to show presence of a very faint band), and no transcript variant 2 was seen. Although a 683 bp band is visible in the PCR analysis for transcript variant 2, had it been present, we would have seen a smaller band of 425 bp (Figure 3. 3a). In this analysis, transcript variant 2 was undetectable in all samples.

Immunofluorescence staining of HaCaT and human primary keratinocytes (HPKs) with anti-TMEM79 rabbit polyclonal antibody show no visible expression of endogenous TMEM79, suggesting that even though the message is present, no TMEM79 protein is being expressed (Figure 3. 3b, c). Controls are shown in Figure 3. 3d and e, where HaCaT cells and HPKs were transfected with 0.2 μg

TMEM79/pEGFPC1 and co-immunofluorescence with anti-TMEM79 antibody shows co-localization with EGFP signal.

The existence of transcript variant 1a and 1b in human epidermis was supported by quantitative RT-PCR of a human multiple tissue array using a predesigned ABI TaqMan® probe for human *TMEM79* (Figure 3. 4), which showed high expression of *TMEM79* in stratified epithelial tissues; cervix, skin and tongue. The highest expression of *TMEM79* was found in prostate tissue. The probe used for *TMEM79* spans across exons 3 and 4, and would amplify all proposed transcript variants of *TMEM79* since they all share this region.

Individual TaqMan® probes were designed to amplify *TMEM79* exon 1 transcript variants. The results for both transcript variants 1a and 1b are combined together in Figure 3. 5. The number of copies for each tissue in Figure 3. 5, where transcript variants 1a and 1b are combined, are noticeably lower when compared to total TMEM79 in Figure 3. 4. An explanation for this difference could be attributed to the difference in design between the custom probes. The probes and/or primers for each assay might have different binding affinities for their respective target sequences. *TMEM79* transcript variant 1b was expressed in prostate, skin, cervix, tongue and mammary gland, reflecting the results of the predesigned assay for total *TMEM79* (Figure 3. 4), whereas transcript variant 1a was more widely expressed across tissues at lower levels. Another reason for the discrepancy in copy numbers seen in both experiments is that there may be other transcript variants, which have not yet been identified.

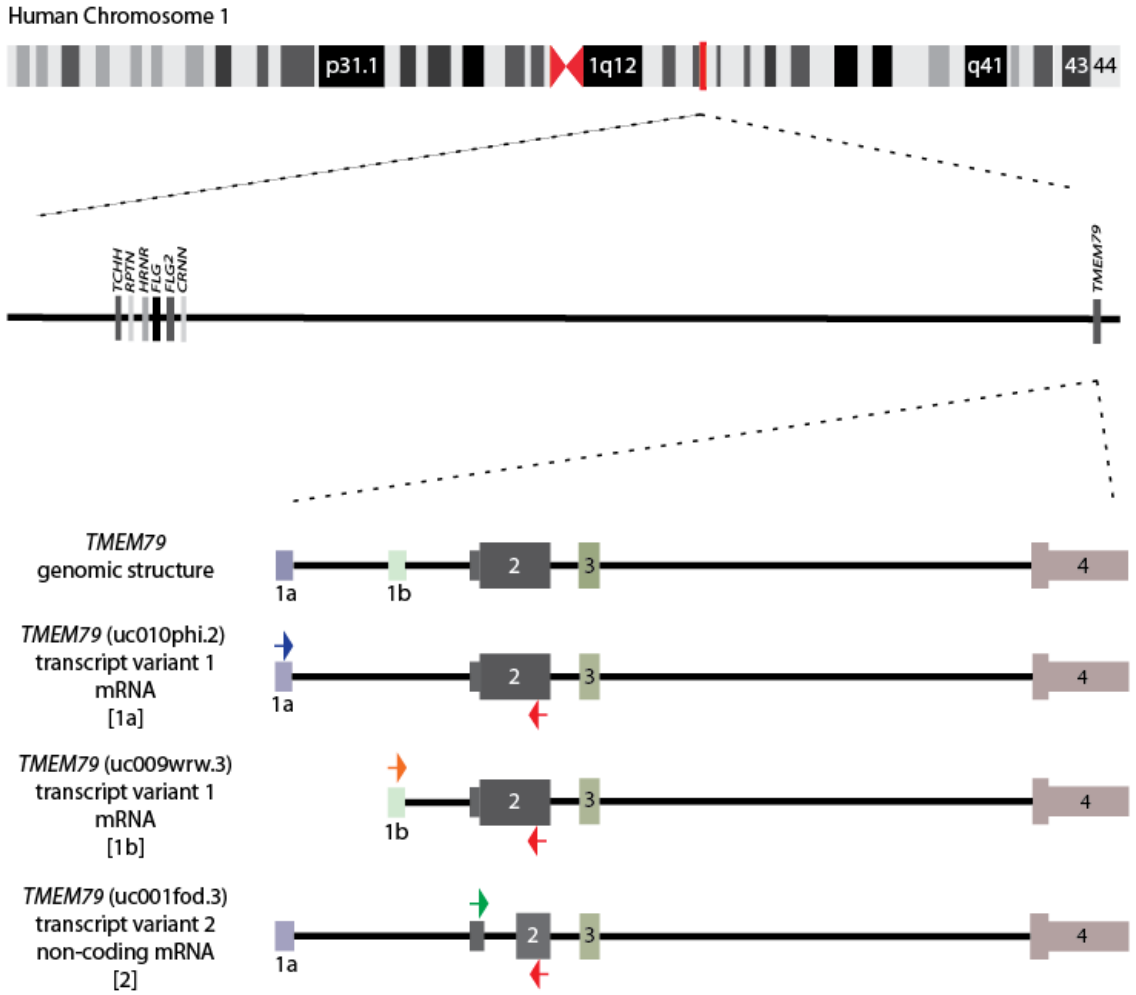



**Figure 3. 2 Genomic location of human TMEM79 with predicted transcript variants.** Location of human ortholog of *TMEM79* on chromosome 1q23.1, showing the location of *TMEM79* in relation to the other EDC genes (Trichohyalin, *TCHH*; Repetin, *RPTN*; Hornerin, *HRNR*; Filaggrin, *FLG*; Filaggrin2, *FLG2*; Cornulin, *CRNN*). Three transcript variants of *TMEM79* show that they all share exon 3 and 4, except for with variations within exon 1 and exon 2. Transcript variant-specific forward primers have been designed for specific amplification of each transcript variant (blue, orange and green arrows) with a common reverse primer (red arrows). The coding regions are indicated by the larger regions, and only start in exon 2.

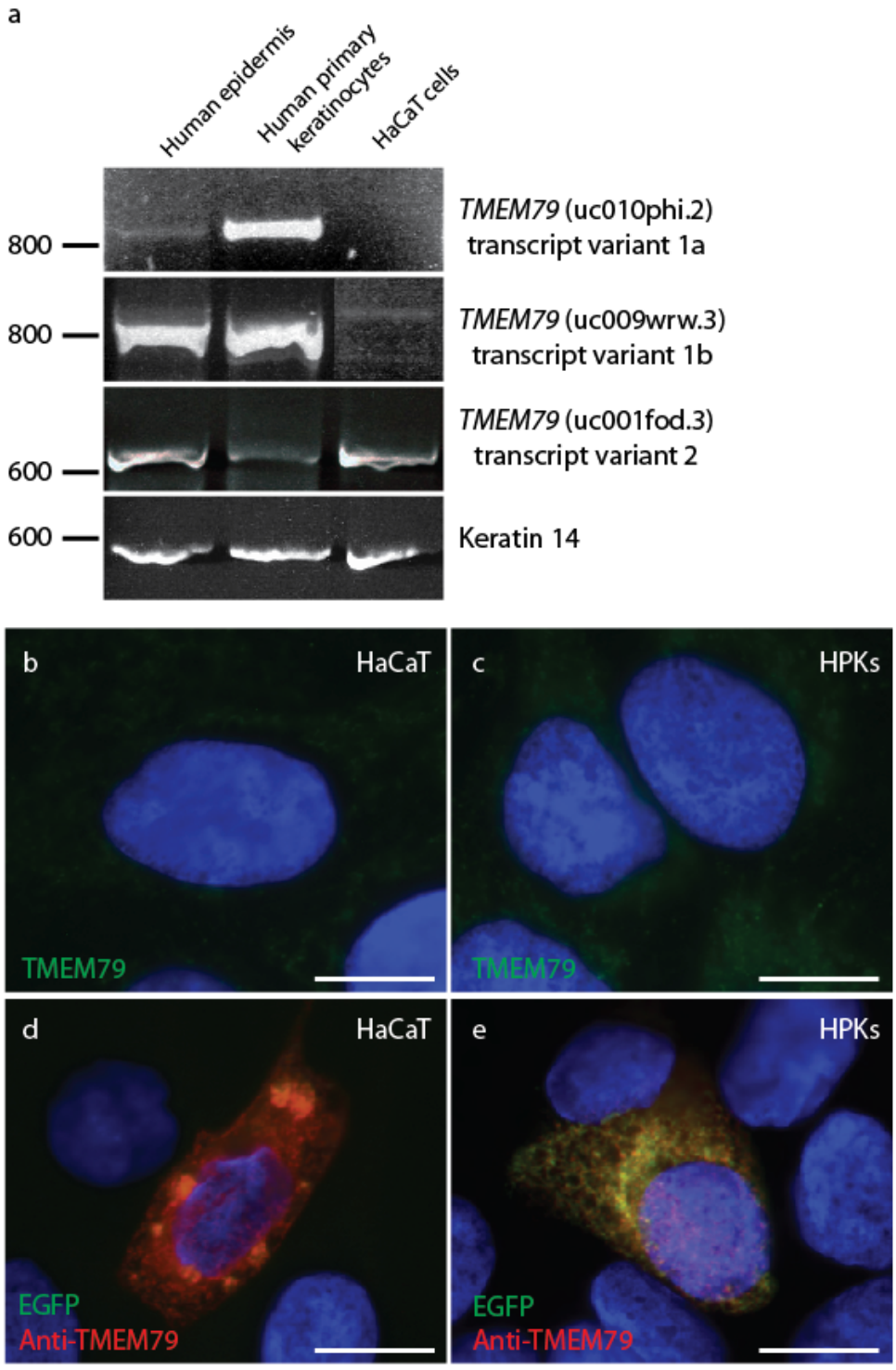



**Figure 3. 3 Semi-quantification of *TMEM79* transcript levels in various cells lines with no detectable endogenous TMEM79.** Semi-quantitative RT-PCR of *TMEM79* cDNA from human epidermis, human primary keratinocytes (HPKs) and HaCaT cells show high quantities of *TMEM79* transcript variant 1b in epidermis and HPKs. No transcript variant 2 was detected (425 bp). Keratin 14 was used as a positive control (a). Immunofluorescent analysis of endogenous TMEM79 show no detectable expression in HaCaT cells (b) or in HPKs (c) with anti-TMEM79 rabbit polyclonal antibody. HaCaT cells transfected with TMEM79/pEGFPC1 and stained with anti-TMEM79 (red) co-fluorescing with EGFP (green) (d and e). Scale bar = 10 µm

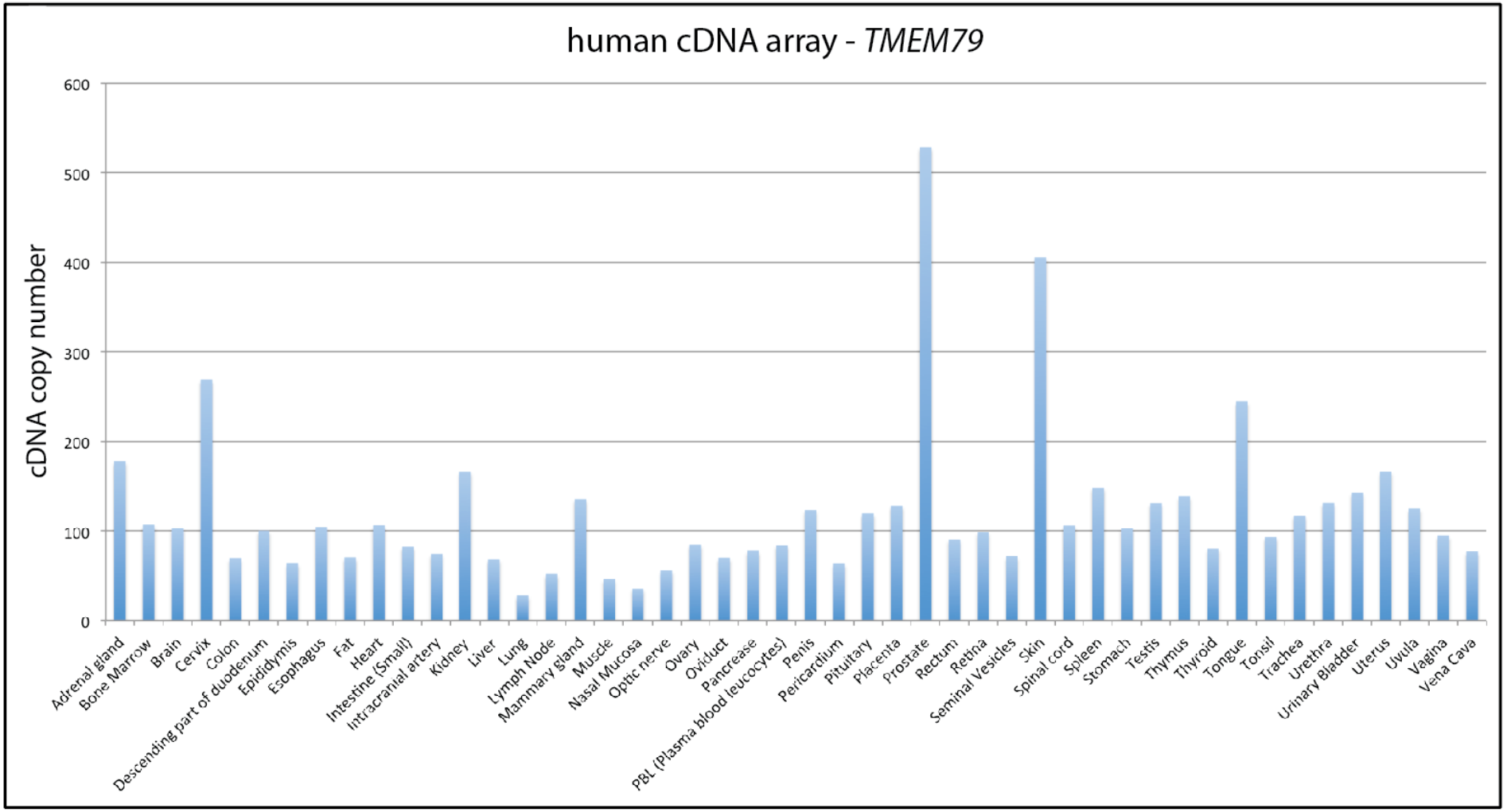



**Figure 3. 4 Multiple tissue expression analysis of total TMEM79.** PCR quantification of human multiple tissue array with probe against total TMEM79 (exon 3-4), showing highest expression in stratified tissues: prostate, skin, cervix and tongue.

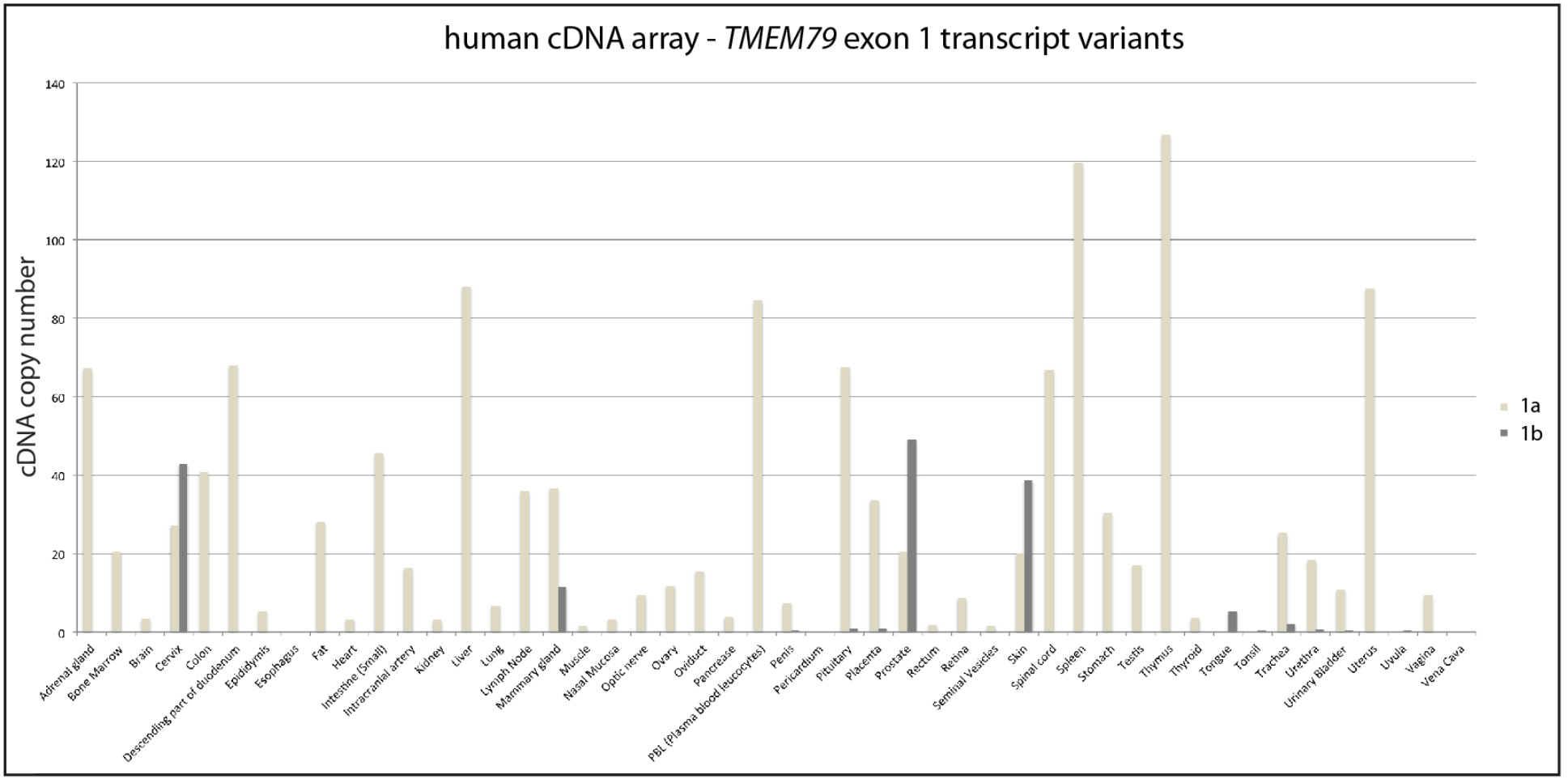



**Figure 3. 5 Multiple tissue expression analysis of *TMEM79* transcript variant 1a and 1b.** PCR quantification of *TMEM79* exon 1 transcript variants in a human multiple tissue array using different probes specific to either transcript variant 1a (beige bars) or 1b (green bars). No specific expression pattern was observed for transcript variant 1a. Highest levels of transcript variant 1b were expressed stratified tissues, which showed highest levels of total *TMEM79* seen in Figure 3. 4.

### 3.3.2. *TMEM79* cDNA sequence similarity searches

In order to investigate if TMEM79 was present in other species apart from human and mice, a sequence similarity search was performed using the human cDNA *TMEM79* sequence (uc010phi.2, transcript variant 1a) from the Uniprotkb database using BLASTX version 2.0MP-WashU (Gish and States, 1993). The results returned showed that the highest identities were found in the apes, chimpanzees, gibbons and gorillas, with close to 90% identity similarity and this included other mammals such as bats, horses, cows, dogs and cats (labeled as "sequence match" in Figure 3. 6). The search also identified transcripts with high similarity to *TMEM79* in zebra fish, turtles and chicken; these transcripts were ranked according to identity match, with red being the most similar and blue the least (Figure 3. 6). However, as the cDNA sequences were not well characterized in most other species, apart from human and mouse, most sequences from 5' and 3' untranslated regions were unavailable. A search for protein sequence similarity was also carried out (labeled as subject match in Figure 3. 6), showing high identity matches for all amino acids of the human TMEM79 protein.

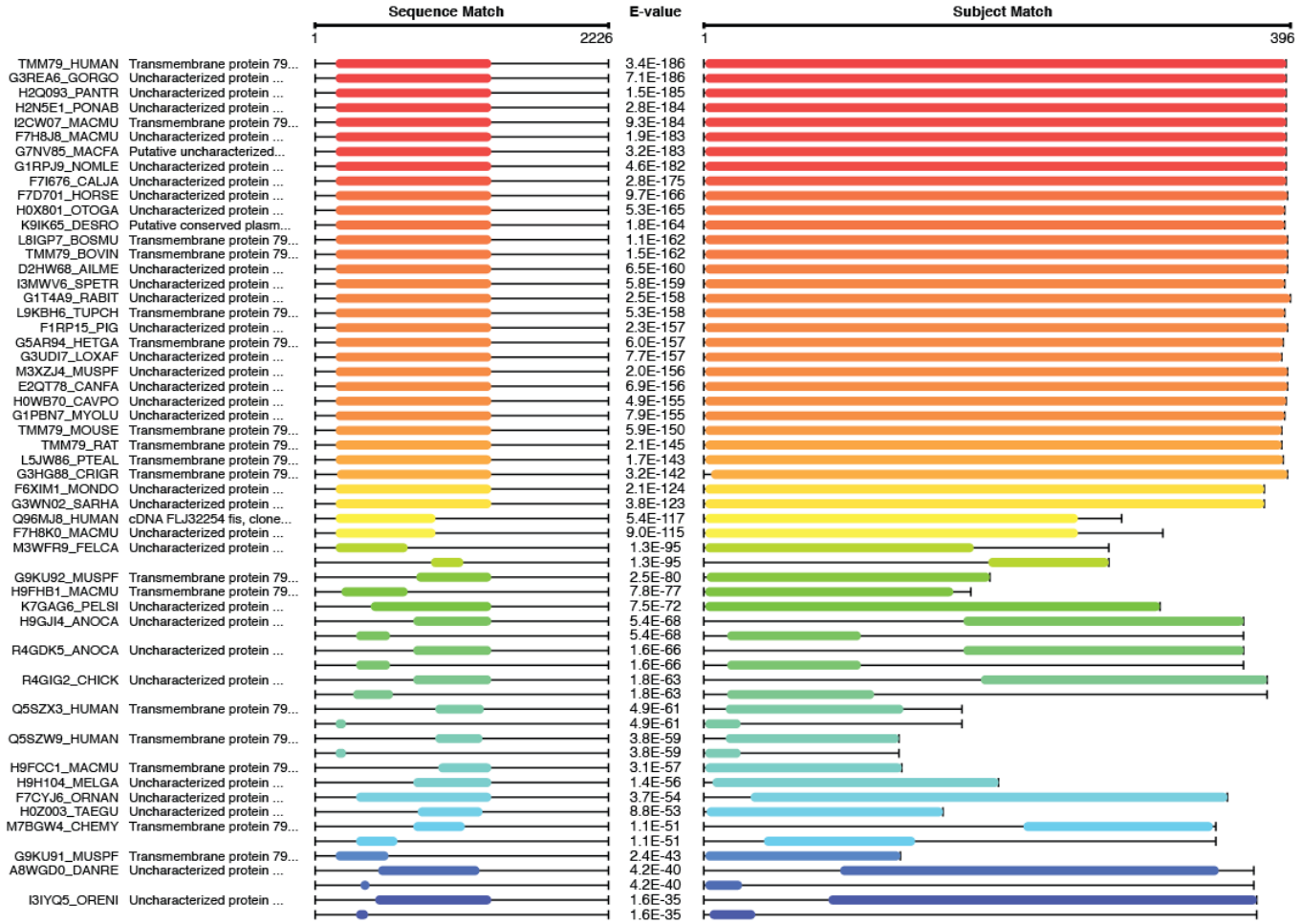



**Figure 3. 6 cDNA sequence alignment search across species using WU-BLAST.** Colour spectrum indicates sequence similarity to human *TMEM79* cDNA, with red being the highest and blue, the lowest. Sequence match refers to *TMEM79* sequence similarity, while subject match refers to TMEM79 protein sequence similarity across species. The alignment shows significant matches across different species showing evolutionary conservation of TMEM79, suggests it might have an important functional role.

### 3.3.3. TMEM79 protein structure prediction

Bioinformatics analysis predicts that the TMEM79 protein consists of a long N-terminal domain, 5 transmembrane domains and a short C-terminus (Figure 3. 7), and this arrangement is hypothesized for both transcript variants of exon 1, since both share identical translational sequences (Figure 3. 2). Protein analysis was carried out by Dr. George Schneider (Bioinformatics Institute, A*STAR, Singapore; Appendix VII) and shows that TMEM79 shares domain homology with members of the MAPEG (Membrane Associated Proteins in Eicosanoid and Glutathione metabolism) superfamily. An alignment with other members of the MAPEG family suggest that the N-terminal domain is located on the inside of a membrane, and the C-terminus on the outside (Figure 3. 7). Although TMEM79 is predicted to have five transmembrane domains, MAPEG members only have four. The transmembrane helices contain proline residues, which are a feature of channel or transporter proteins, and TMEM79 protein also contains key residues at amino acid positions 66, 70 and 126, which are predicted to be involved in glutathione binding.

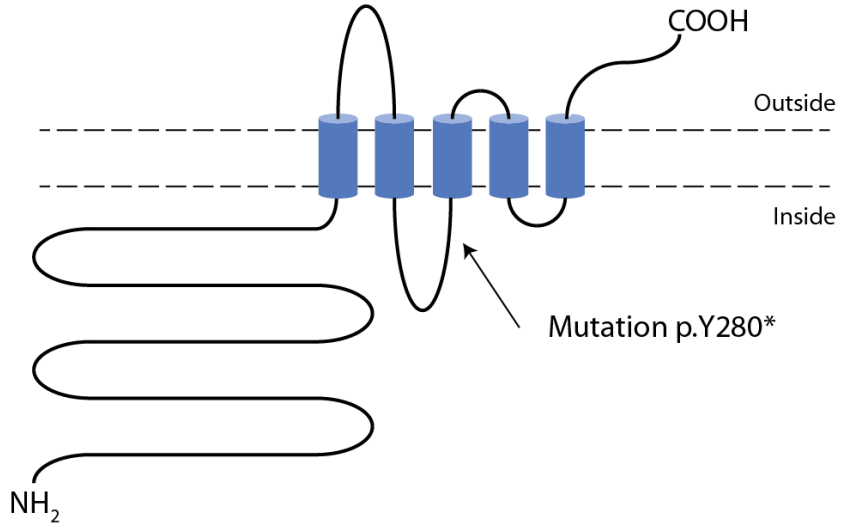



**Figure 3. 7 A cartoon of proposed secondary structure for TMEM79.**
Bioinformatics analysis predicts 5 transmembrane domains, with a long N-terminal domain and a short external C-terminal domain. Arrow shows the location of the matted mutation, p.Y280*, on the first internal loop.

Pairwise alignment between human and mouse TMEM79 protein sequences show that they share 83% identity (SIM, ExPaSy Bioinformatics Resource Portal; Figure 3. 8). Murine TMEM79 protein has 3 amino acids less than human TMEM79 protein, and this distinction is located within the N-terminal region. Due to the close homology of both proteins, it can be assumed that both proteins would have very similar functions in their respective hosts.

```
83.0% identity in 394 residues overlap; Score: 1702.0; Gap frequency: 0.8%

Human          1 MTEQETLALLEVKRSDSPEKSSPQALVPNGRQPEGEGGAESPGAESLRVGSSAGSPTAIE
Mouse          1 MTEPETLALLDMKEPETPEKSPPQALV---LQSEEEGGTESPGTESLRVGSSVGSPIVRE
                 *** ******  *    **** *****    * * *** **** ******** ***   *

Human         61 GAEDGLDSTVSEAATLPWGTGPQPSAPFPDPPGWRDIEPEPPESEPLTKLEELPEDDANL
Mouse         58 GPEDGPDSTISEAATLPWGTDPHPSAPLPDPPGWRDIEPEPLESEAPTKSEEPFKEDANL
                 * *** *** ********** * **** ************* ***  ** **    ****

Human        121 LPEKAARAFVPIDLQCIERQPQEDLIVRCEAGEGECRTFMPPRVTHPDPTERKWAEAVVR
Mouse        118 LPEKTVRAFVPIDLQCIERKPQEERILHRDAGPGELRNFLPARLSHPEPPERKWAEAVVR
                 ****  ************* ***  *    ** ** * * * *  ** * **********

Human        181 PPGCSCGGCGSCGDREWLRAVASVGAALILFPCLLYGAYAFLPFDVPRLPTMSSRLIYTL
Mouse        178 PPGRSCGGCGSCGGREALRAVASVVAALIFFPCLLYGAYAFLPFDAPRLPTMSSRLVYTL
                 *** ********* ** ******* **** ************** ********** ***

Human        241 RCGVFATFPIVLGILVYGLSLLCFSALRPFGEPRREVEIHRRYVAQSVQLFILYFFNLAV
Mouse        238 RCGVFATFPIVLGLLVYGLSLLCFSALRPFGEPRREVEIHRQYVAQSVQLFILYFFNLAV
                 ************* ************************** ******************

Human        301 LSTYLPQDTLKLLPLLTGLFAVSRLIYWLTFAVGRSFRGFGYGLTFLPLLSMLMWNLYYM
Mouse        298 LSTYLPQDTLKLLPLLTGLFAISRLIYWLTFAVGRSFRGFGYGLTFLPLLAMLVWNLYYM
                 ******************** ***************************** ** ******

Human        361 FVVEPERMLTATESRLDYPDHARSASDYRPRPWG
Mouse        358 FVVEPERMLTASESRLDYPDHARSVSDYRPRSWG
                 *********** ************ ****** **
```

**Figure 3. 8 Pairwise alignment between human and murine TMEM79 protein sequences.** ' * ' indicates identical residues. With 83% identity match, this suggests that both human and murine TMEM79 proteins would have similar functions.

Using the human TMEM79 protein sequence, a multiple sequence alignment with TMEM79 sequences from naked mole rat, chicken and zebrafish was carried out using the Clustal Omega program from The European Bioinformatics Institute (EBI) (Goujon et al., 2010; Sievers et al., 2011). The N-terminal region of the protein is poorly conserved, but there are a high number of conserved residues seen within the predicted transmembrane domains (numbered red boxes). Of

particular interest, are the residues in between the transmembrane regions and at the C terminus (underlined in green), suggesting that these evolutionarily-preserved regions probably contribute to functional importance (Figure 3. 9), as they are all in locations appropriate for protein-protein interaction.

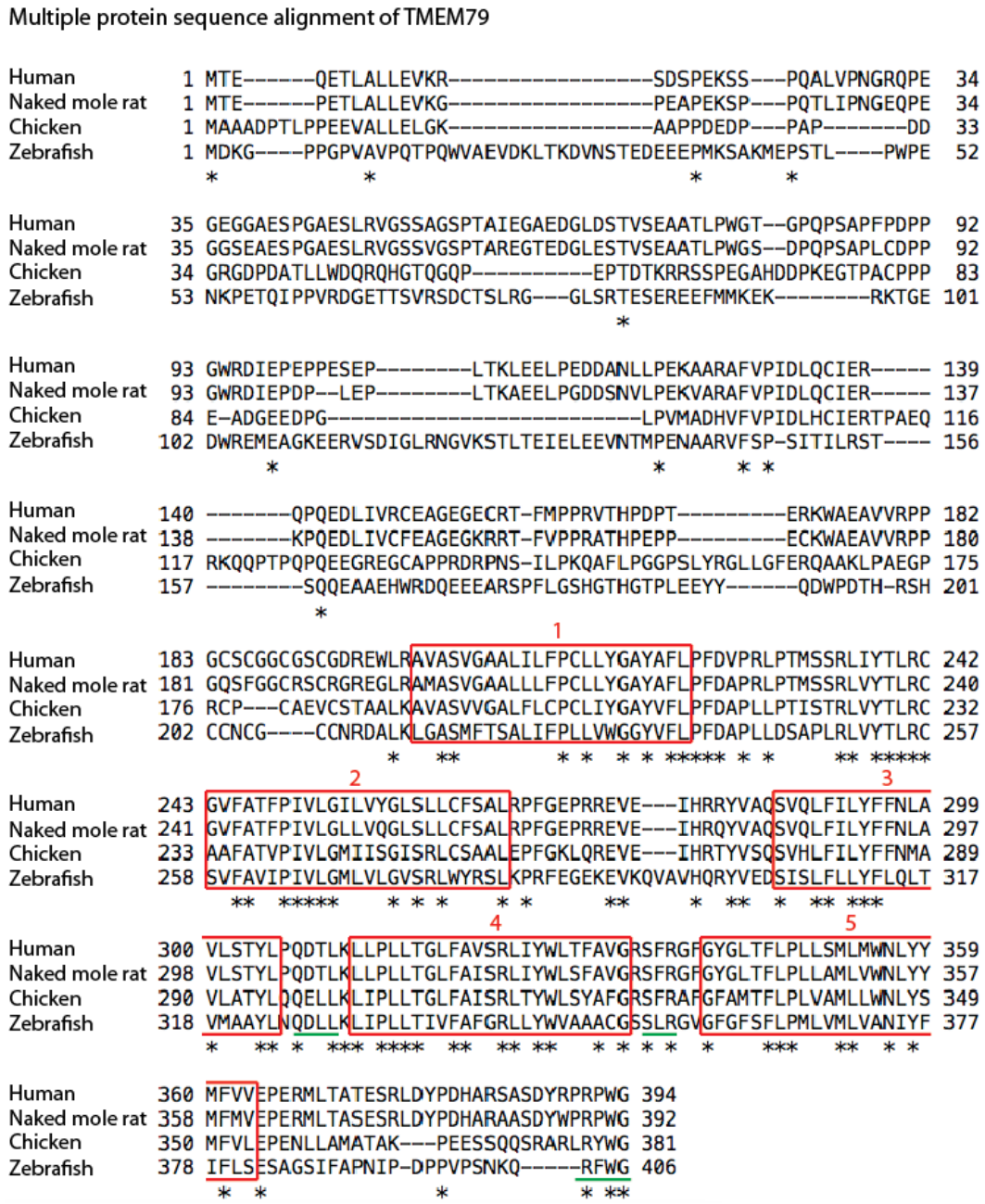



**Figure 3. 9 Protein sequence alignment of TMEM79 in humans, naked mole rat, chicken and zebrafish.** Numbered red boxes indicate transmembrane domains, and show remarkable identity match across species. Residues at the transmembrane domain and C-terminal region showing conservation across species. ' * ' indicates identical residues, ' : ' and ' . ' indicates partial matches, and green underlined residues show remarkably high conservation at the C-terminus end of TMEM79.

### 3.3.4. Localization of human TMEM79 tissue expression

Immunofluorescence microscopy showed that the human ortholog of TMEM79 was expressed in the upper granular layer of the epidermis (Figure 3. 10), with strong staining in the hair follicle (Figure 3. 10a), similar to the staining seen in mouse epidermis. The stronger staining seen in human skin compared to mouse skin could be attributed to the general thickness of the epidermis, compared to murine neonatal skin. Another reason could be antibody affinity, as the antibody used was raised against a human TMEM79 sequence. Brightfield illumination in combination with immunofluorescence shows that the presence of TMEM79 can clearly be seen in the layer beneath the stratum corneum (seen in red, Figure 3. 10c, d).

An antibody to keratin 10 was used as a marker for the suprabasal layers (Figure 3. 11a), and although the antibody showed some non-specific reactivity to the SC, there is clear cytoplasmic staining within the granular layers, and keratin 10 shows partial co-localization with TMEM79 within the upper layers. The basal cell layer was stained with a keratin 14 antibody (Figure 3. 11b). Co-immunofluorescence with desmosomal junctional markers desmoglein 1 (DSG1; Figure 3. 11c) and desmoplakin 1/2 (DSP1/2; Figure 3. 11d) show that TMEM79 expression in the upper granular layers shares a similar junctional staining pattern. Corneodesmosin (CDSN) showed a cytoplasmic staining pattern in the upper granular layers, before migrating to the corneodesmosome junctions between corneocytes (Figure 3. 12a). Kallikrein 8 (KLK8) is expressed at the topmost granular layers (Figure 3. 12b) and both LG-secreted proteins show localization to the same upper granular layers as TMEM79. Stratum corneum markers such as filaggrin (FLG), involucrin (INV) and loricrin (LOR) show no colocalization with TMEM79 (Figure 3. 12c - e). This confirms that TMEM79 is

located at the interface between the SG and SC as keratinocytes enter terminal differentiation from the granular to cornified layers.

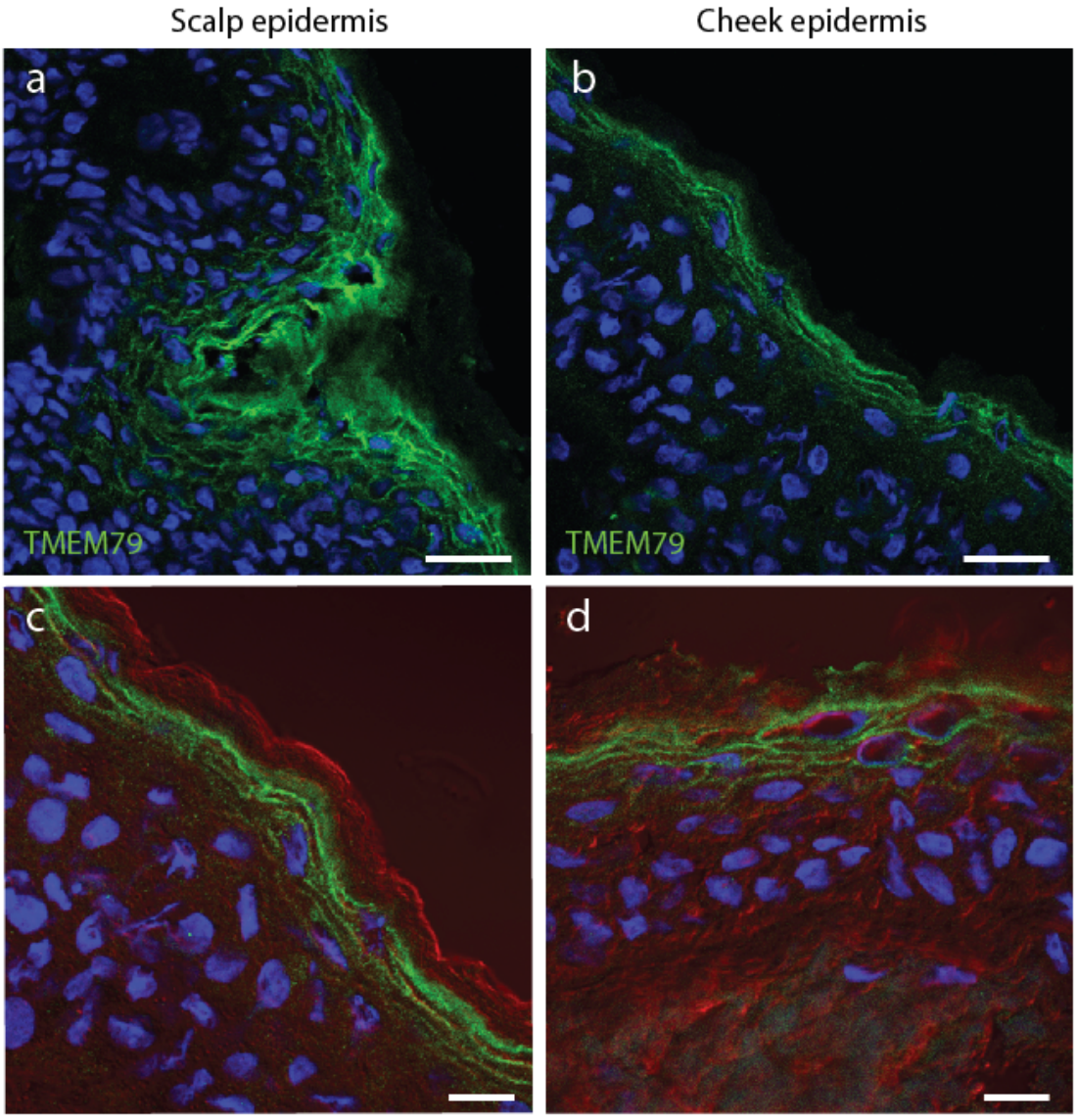



**Figure 3. 10 Immunofluorescent analysis of TMEM79 expression pattern in human skin sections.** Sections of scalp and cheek epidermis (as indicated) were stained with a rabbit polyclonal antibody against TMEM79 (green) and show expression within the upper granular layers. Stained sections are shown illuminated with bright field (red) to show that TMEM79 expression does not extend into the layers of the stratum corneum (c, d). Nuclei were stained with DAPI (blue). Scale bar = 20 μm

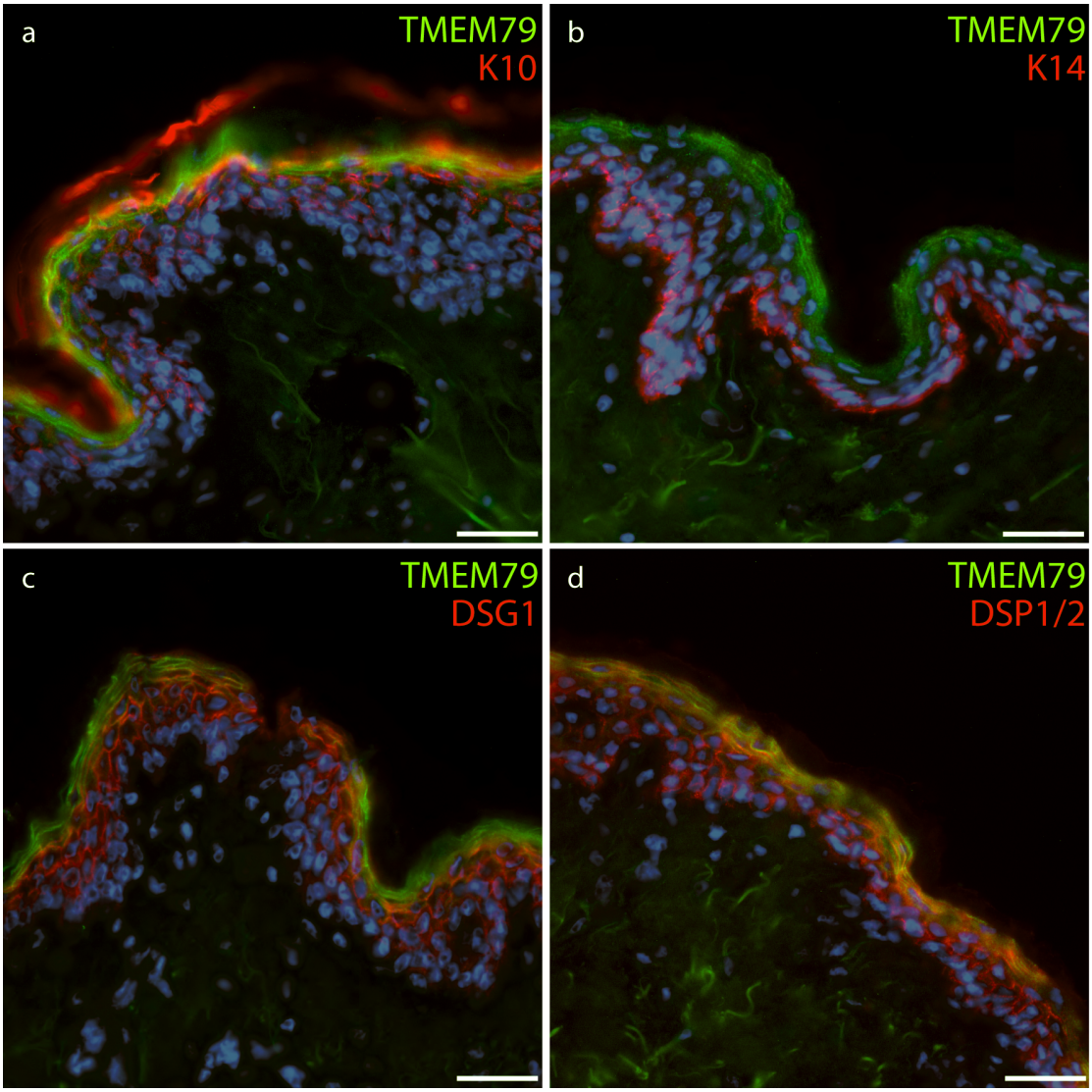



**Figure 3. 11 Co-immunofluorescent analysis of TMEM79 in human abdominal skin sections with epidermal markers.** Co-localization with keratin 10 is seen in the suprabasal layers (a), but not with keratin 14 in the basal layer (b). Epidermal junction markers for desmosomes show partial co-localization with desmoplakin 1/2 (DSP1/2) (d), and desmoglein 1 (DSG1) (c) and similar junctional staining. Nuclei were stained with DAPI (blue). Scale bar = 50 μm

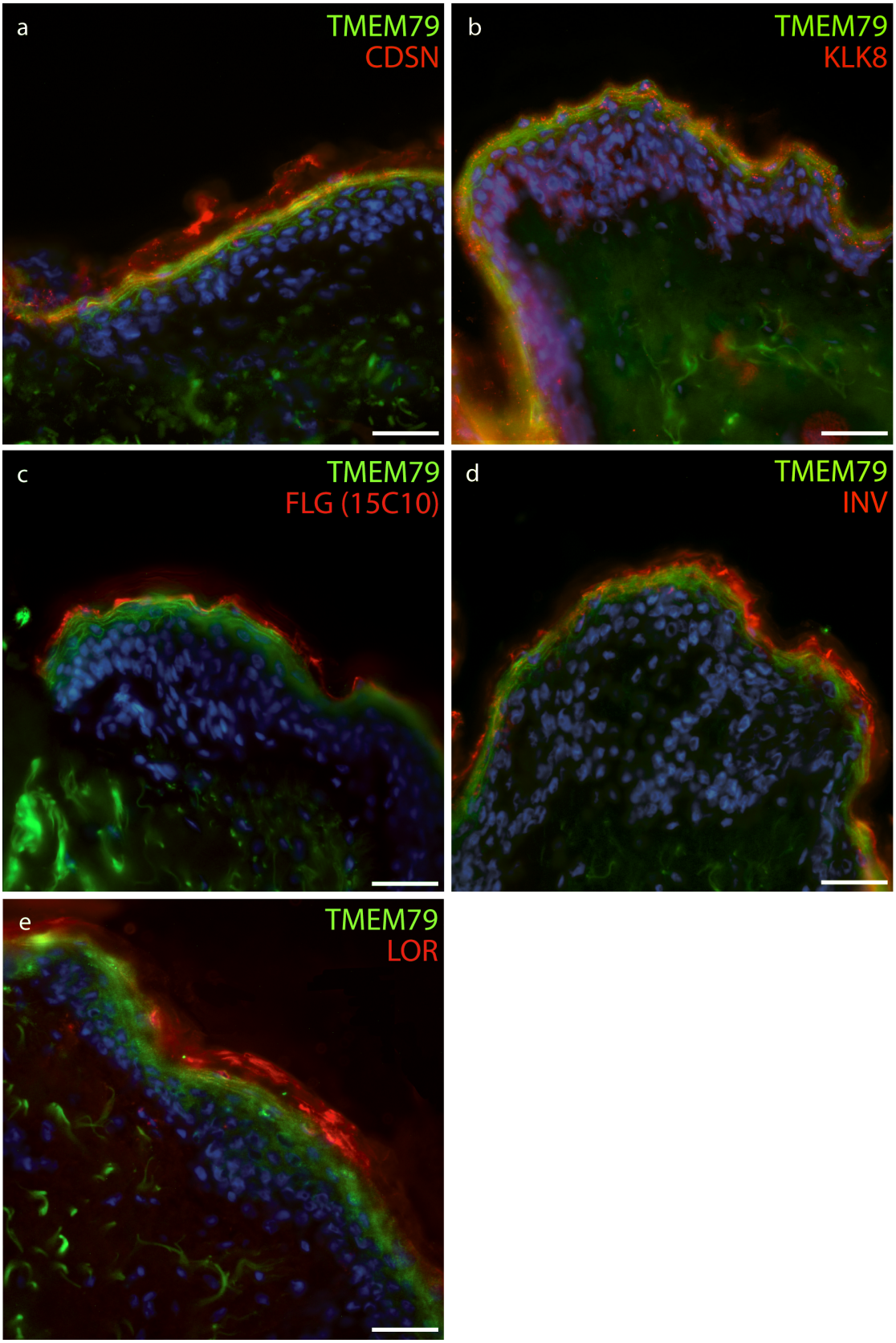



**Figure 3. 12 Co-immunofluorescent analysis of TMEM79 in human abdominal skin sections with terminal differentiation markers.** Lamellar granule secreted proteins corneodesmosin (CDSN) and kallikrein 8 (KLK8) show localization to the same layers as TMEM79 (a, b). No co-localization was seen with stratum corneum markers; filaggrin (FLG) (c), involucrin (INV) (d) and loricrin (LOR) (e). Nuclei were stained with DAPI (blue). Scale bar = 50 µm

### 3.3.5. Hair follicle expression

As immunofluorescence of epidermis shows strong expression of TMEM79 within the HF, further investigation of the human hair follicle using immunofluorescence confirmed that TMEM79 is located in the inner root sheath in both longitudinal (Figure 3. 13a, c) and cross-sections of hair follicles (Figure 3. 13b), in contrast to K17, a marker for the outer root sheath (ORS) of the hair follicle (Figure 3. 13; red). TMEM79 appears to have a junctional, possibly punctate staining pattern upon examination at a higher magnification (Figure 3. 13d, e). Specific markers for the inner root sheath revealed that TMEM79 co-localized with keratin 74, which is specific for the Huxley layer (Figure 3.14, Panel B). This staining did not extend to the keratin 32 expression which is specific for the inner hair cuticle layer (Figure 3.14; Panel A), or to the keratin 75-specific outer companion layer (Figure 3.14; Panel C). It is possible that the Henley layer might stain positive, but there isn't a specific antibody available for that layer.

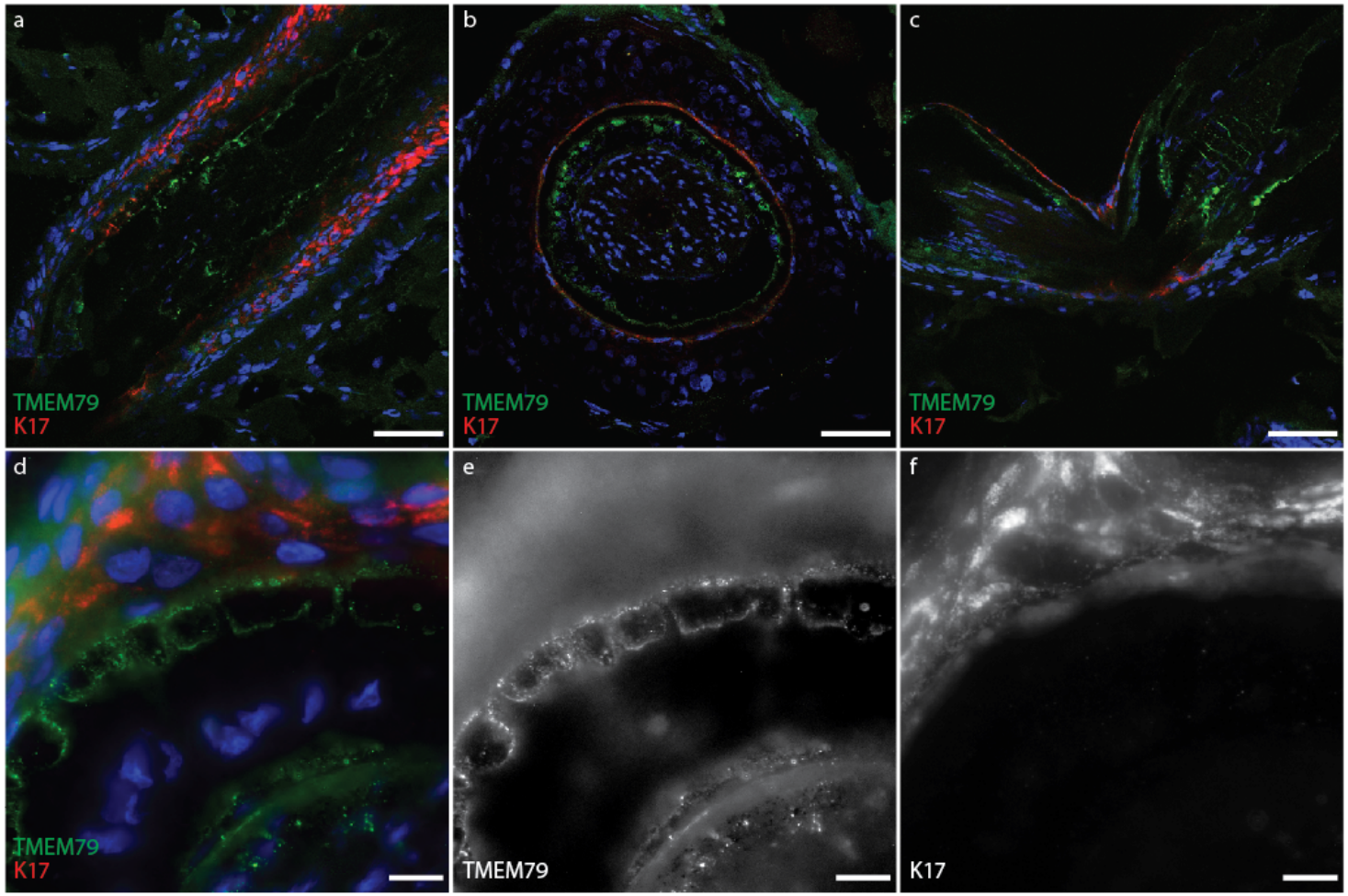



**Figure 3. 13 Co-immunofluorescent analysis of TMEM79 in human hair follicle sections with keratin 17.** An outer root sheath (ORS) marker, keratin 17 (red), shows no co-localization with TMEM79 (green); longitudinal (a, c) and cross-section (b) views. At higher magnification, TMEM79 within the IRS shows a punctate, possibly junctional staining of TMEM79 (d, e) and clearly shows no overlap with K17 in the ORS (f). Nuclei were stained with DAPI (blue). Scale bars for a–c = 50 μm, d–f = 10 μm.

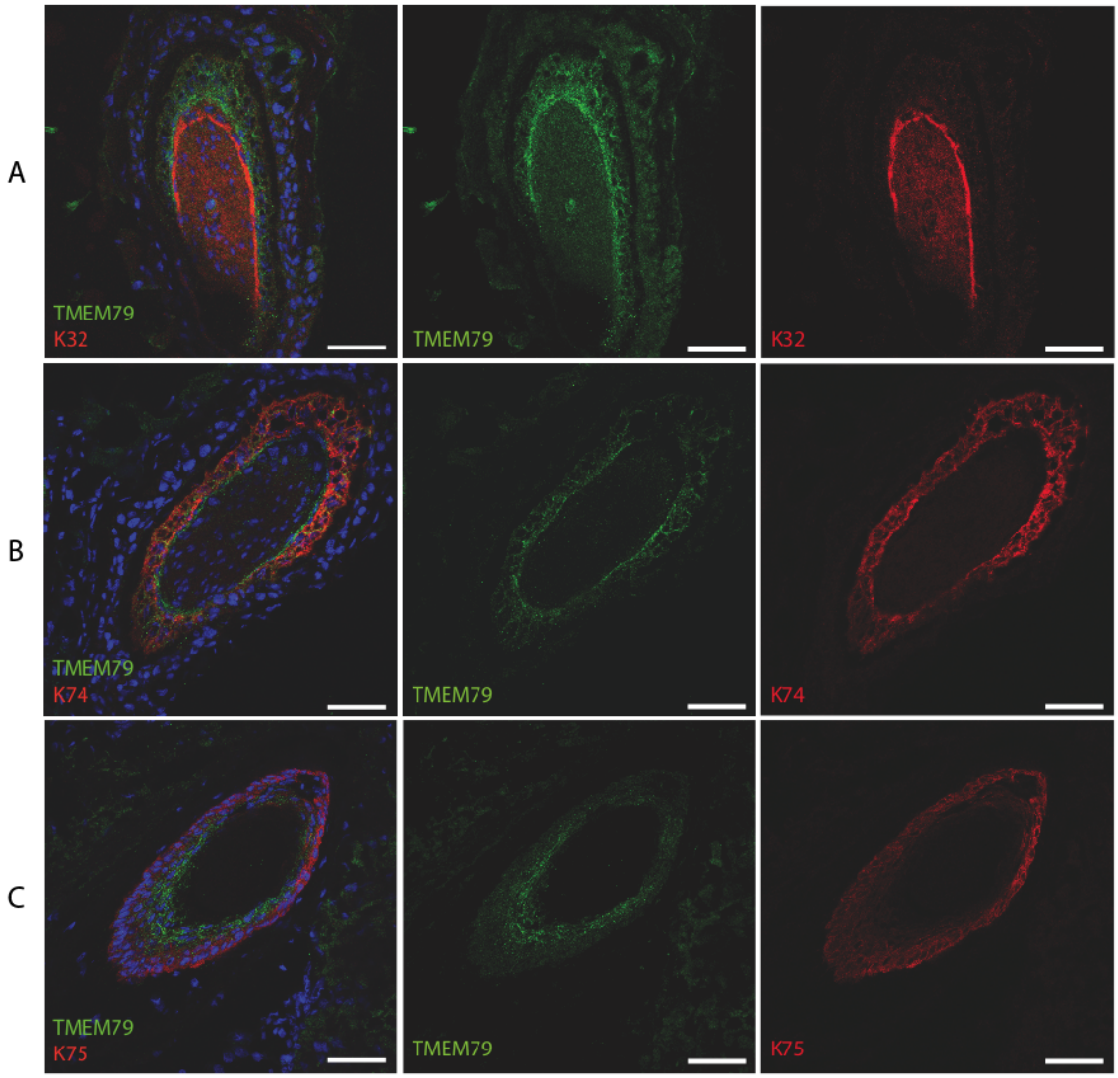



**Figure 3. 14 Immunofluorescent analysis of TMEM79 in human hair follicle sections with inner root sheath markers.** Co-localization of TMEM79 is seen with the Huxley layer (keratin 74; Panel B). TMEM79 staining doesn't extend beyond the Huxley layer, and show no co-localization with the outer companion layer (keratin 75; Panel C) or the inner hair cuticle (keratin 32; Panel A). Leftmost column shows merged images for each immunofluorescent staining. Nuclei were stained with DAPI (blue). Scale bar = 50 μm

### 3.3.6. Immuno-gold labeling of TMEM79 in the human epidermis

To look at the subcellular location of TMEM79, immuno-gold labeling of ultra-thin sections of human abdominal epidermis was performed. Localization of TMEM79 protein to the membranes of lamellar granules (LG) between the topmost granular layer (labeled as SG1 in Figure 3. 15) and the lowest stratum corneum layer (labeled as SC1 in Figure 3.15) was seen. TMEM79 is seen on the membranes of these LGs, but is not present in every LG (Figure 3. 15 a, b; black arrows). These are more clearly seen in the pictures at a higher magnification (Figure 3. 15; higher magnification of the regions within the boxed areas of (a) and (b) are shown in (c) and (d) respectively). At the desmosomal junctions between SG1 and SG2, TMEM79 is seen at a short distance adjacent to desmosomal structures, which could indicate a form of storage (Figure 3. 15a; white arrows). The localization of TMEM79 was uniform across 80 ultra-thin epidermal sections.

Double immuno-gold labeling with TMEM79 (10 nm) and corneodesmosin (CDSN; 15 nm) shows that the TMEM79-labelled LGs do not contain CDSN (Figure 3. 16) and that the two proteins do not associate with each other. Figure 3. 16b shows a LG in the process of fusing with the plasma membrane, and Figure 3. 16c shows TMEM79 at the SC-SG intercellular space after the contents have been emptied out.

Localization of TMEM79 and kallikrein 8 (KLK8) show that they are both associated with LG at the SG-SC interface (Figure 3. 17). LG containing KLK8 are situated further from the SG-SC interface, with the contents from TMEM79-labelled granules being emptied out first (Figure 3. 17.c, d).

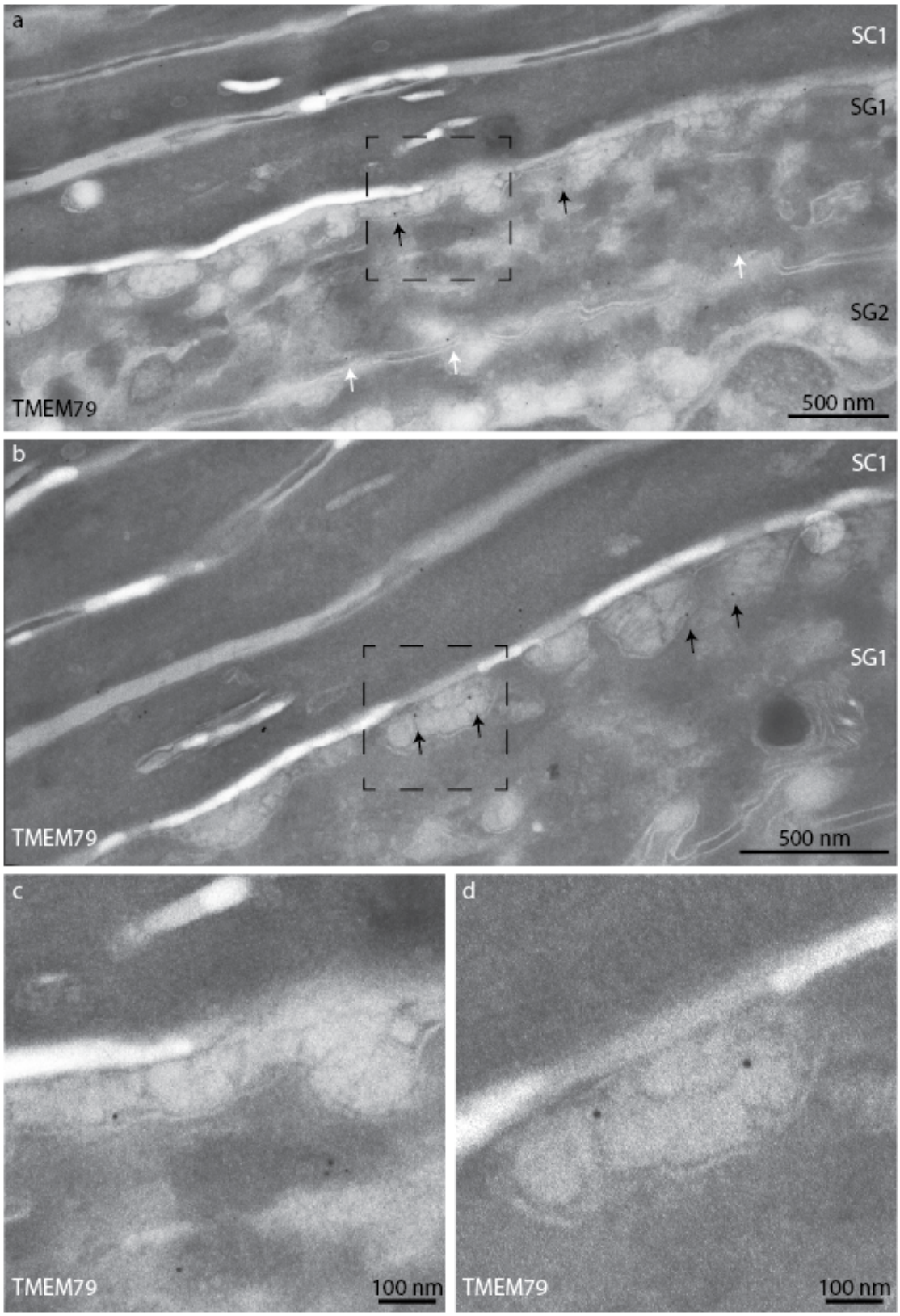



**Figure 3. 15 Immuno-gold TMEM79 labeling of human abdominal epidermis.** TMEM79 is seen located within the membrane of lamellar granules between the top stratum granular layer (SG1 and SG2) and stratum corneum (SC; a and b, black arrows). Between SG1 and SG2, there is gold labeling adjacent to desmosomal structures, indicating a possible location for storage of TMEM79 until it is required (a; white arrows). Higher magnification is shown; (c) for boxed area of (a) and (d) for boxed area of (b). Localization of TMEM79 was uniform across a total of 80 ultra-thin sections.

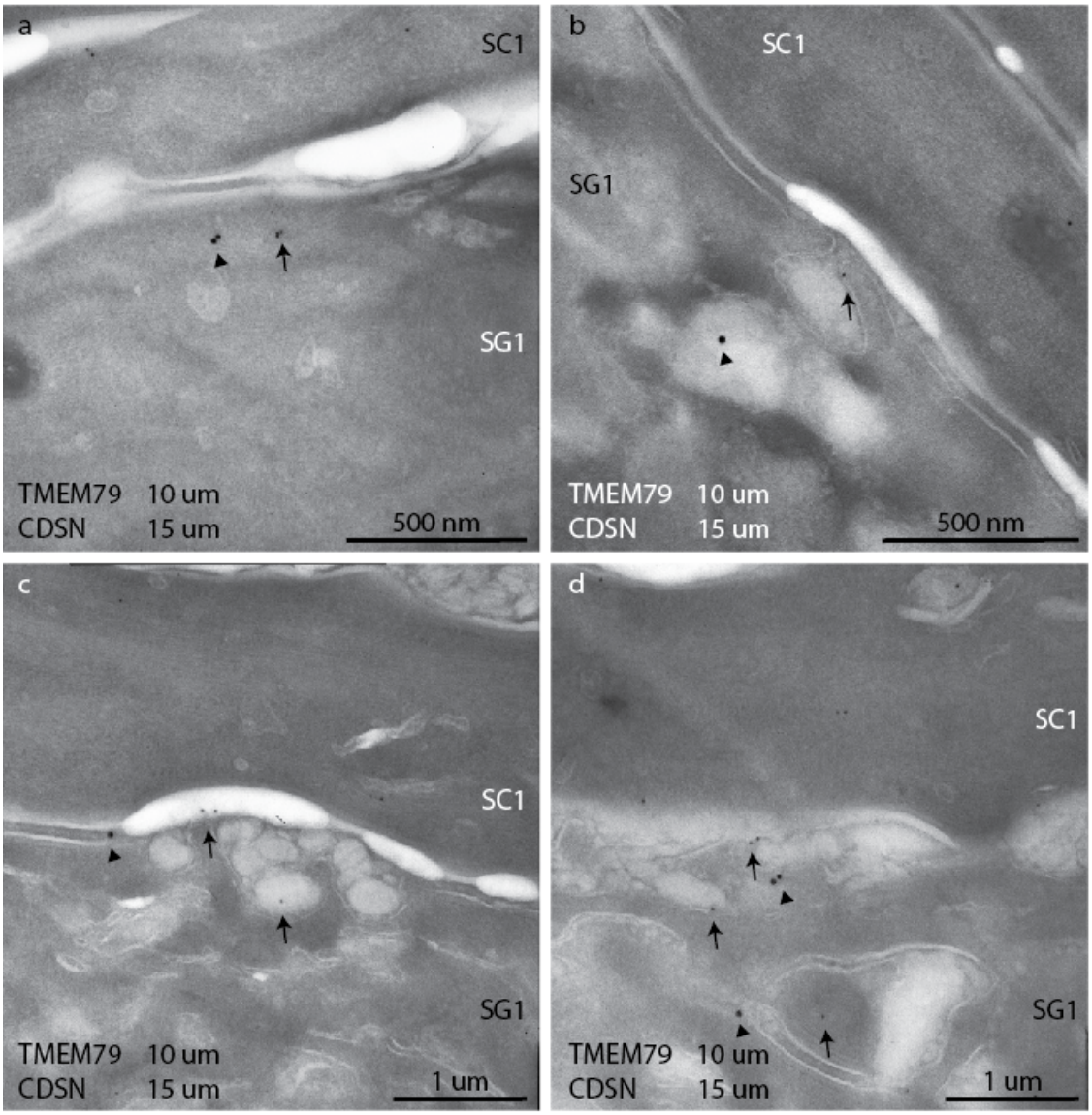



**Figure 3. 16 Double immuno-gold TMEM79 labeling of human abdominal epidermis with corneodesmosin.** No-association was seen between TMEM79 (10 nm gold particles, arrows) and corneodesmosin (CDSN; 15nm gold particles, arrowheads) shows that these two proteins are stored in different lamellar granules. SC1, stratum corneum; SG1, stratum granulosum.

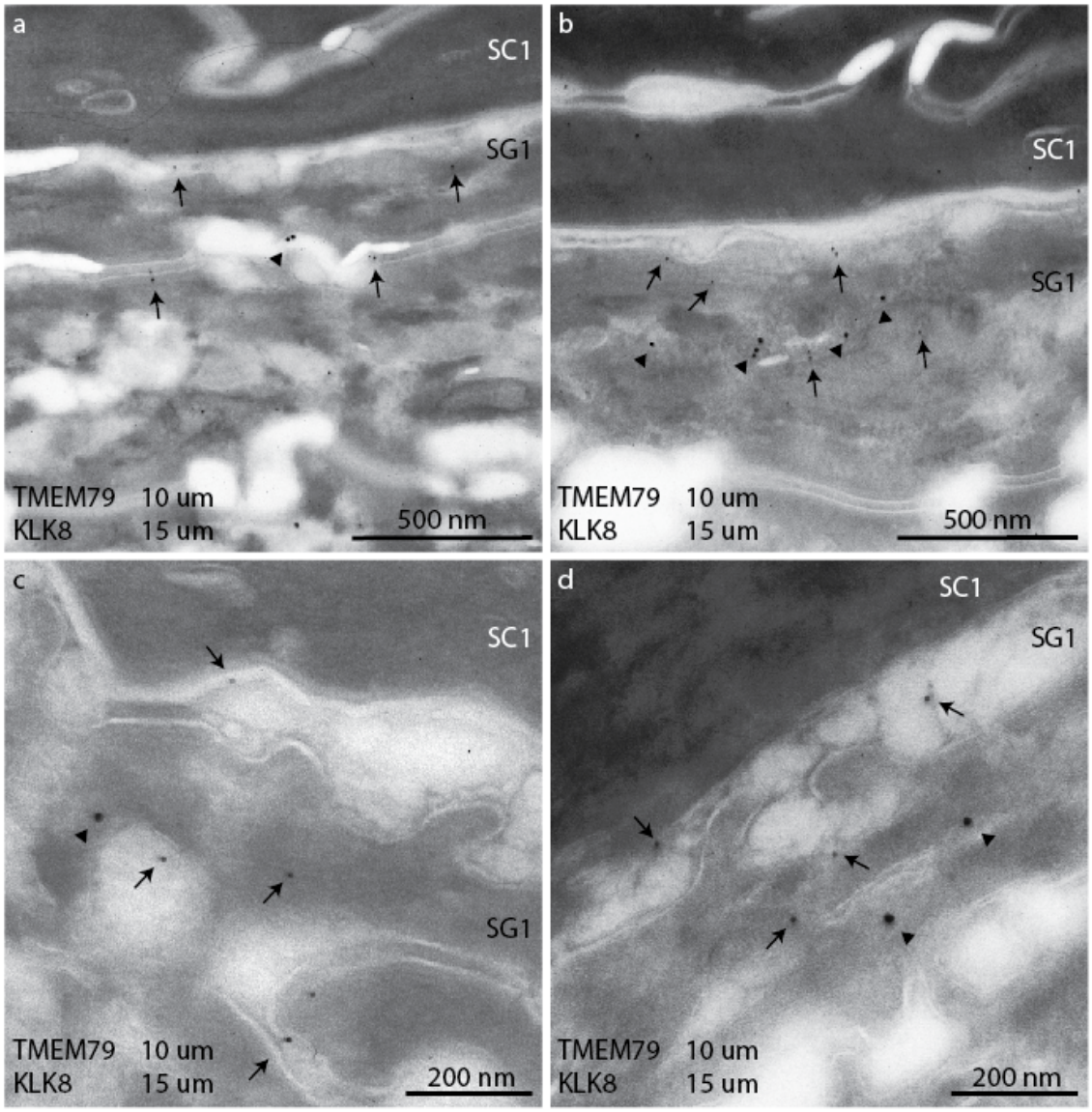



**Figure 3. 17 Double immuno-gold TMEM79 labeling of human abdominal epidermis with KLK8.** TMEM79 (10 nm gold particles; arrows) and KLK8 (15 nm gold particles; arrowheads) shows no association of TMEM79 with KLK8. SC1, stratum corneum; SG1, stratum granulosum.

### 3.3.7. Detection of TMEM79 from human skin and HaCaT cells

To ensure that transfections with TMEM79 constructs were performed correctly, extractions of total protein from normal human abdominal skin (ID9179) and HaCaT cells were analyzed for the presence of TMEM79 protein. The resulting total protein extract from human skin was very viscous; therefore two different dilutions for human skin extracts were loaded. A diffuse band was seen at approximately the predicted size of ~39 kDA for protein extracts from human skin, but not in HaCaT cells. TMEM79 might undergo post-translational modification, which may alter its size and explain the other non-specific bands. The undetectable levels of TMEM79 in HaCaT cells were expected as very low levels of *TMEM79* message were previously detected by semi-quantitative RT-PCR (Figure 3. 3a, page 91). The bands seen for HaCaT cell total protein extracts could be due to antibody non-specific binding or unknown isoforms of TMEM79.

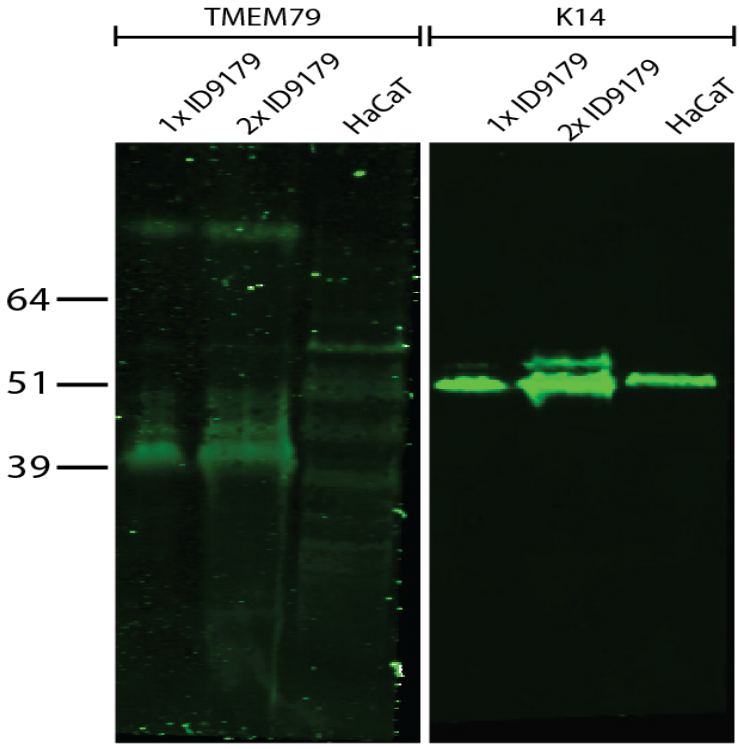



**Figure 3. 18 Western blot analysis of TMEM79 in human skin and HaCaT cells.** Total protein extracts were analyzed by SDS-PAGE and a band was detected at the predicted size of ~39 kDa using an anti-TMEM79 rabbit polyclonal antibody. An antibody to keratin 14 was used as a control.

### 3.3.8. Transient overexpression of TMEM79 constructs in HaCaT cells

#### 3.3.8.1. Myc-TMEM79 overexpression

In an attempt to establish an expression system, HaCaT cells were transfected with a commercial construct, myc-TMEM79, and incubated for 24 hours, following which cells were fixed and analyzed by immunofluorescence using an anti-myc mouse monoclonal antibody. Two different localization patterns were observed (Figure 3. 20). At the lowest transfection concentration of 0.5 µg plasmid DNA per well of a 6-well dish, small aggregates could be seen in a vesicle-like pattern, as well as a network-pattern of expression. At concentrations above 0.5 µg per well, these small aggregates could no longer be seen, and transfected cells show a general distribution network of protein, in a punctate manner throughout the cell. Cells were also transfected and incubated for 48 and 96 hours prior to fixing and staining. Both incubation periods showed no distinctive expression of TMEM79 (Figure 3. 21), suggesting that either the exogenous protein had been degraded, or that over-expression of myc-TMEM79 was lethal to transfected cells over that time period. There was no expression seen in untransfected controls for 24, 48 and 96 hours stained with anti-myc mouse monoclonal antibody (Figure 3. 19).

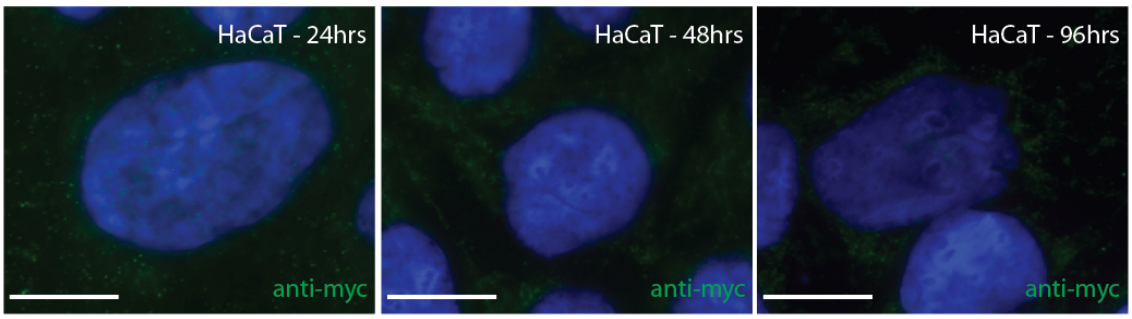



**Figure 3. 19 Untransfected control HaCaT cells after 24, 48 and 96 hours.** Cells were stained with mouse monoclonal anti-myc antibody (green) and showed no detectable fluorescence. Nuclei were stained with DAPI (blue). Scale bar = 10 µm

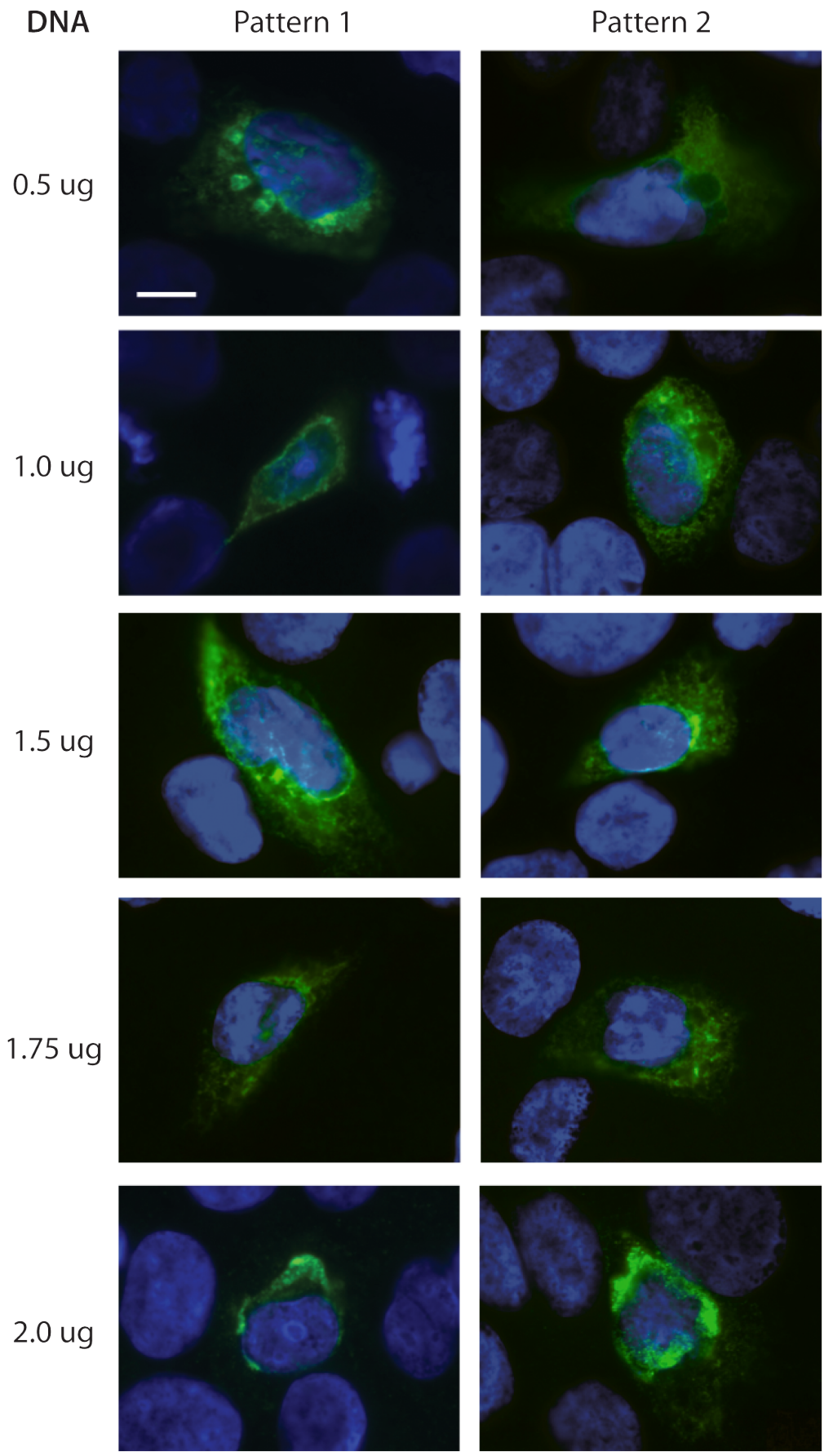



**Figure 3. 20 Immunofluorescent analysis of HaCaT cells transfected with myc-TMEM79 and incubated for 24 hours.** Cells showed two expression patterns after staining with an anti-myc mouse monoclonal antibody (green). Nuclei are stained with DAPI (blue). Scale bar = 10 μm

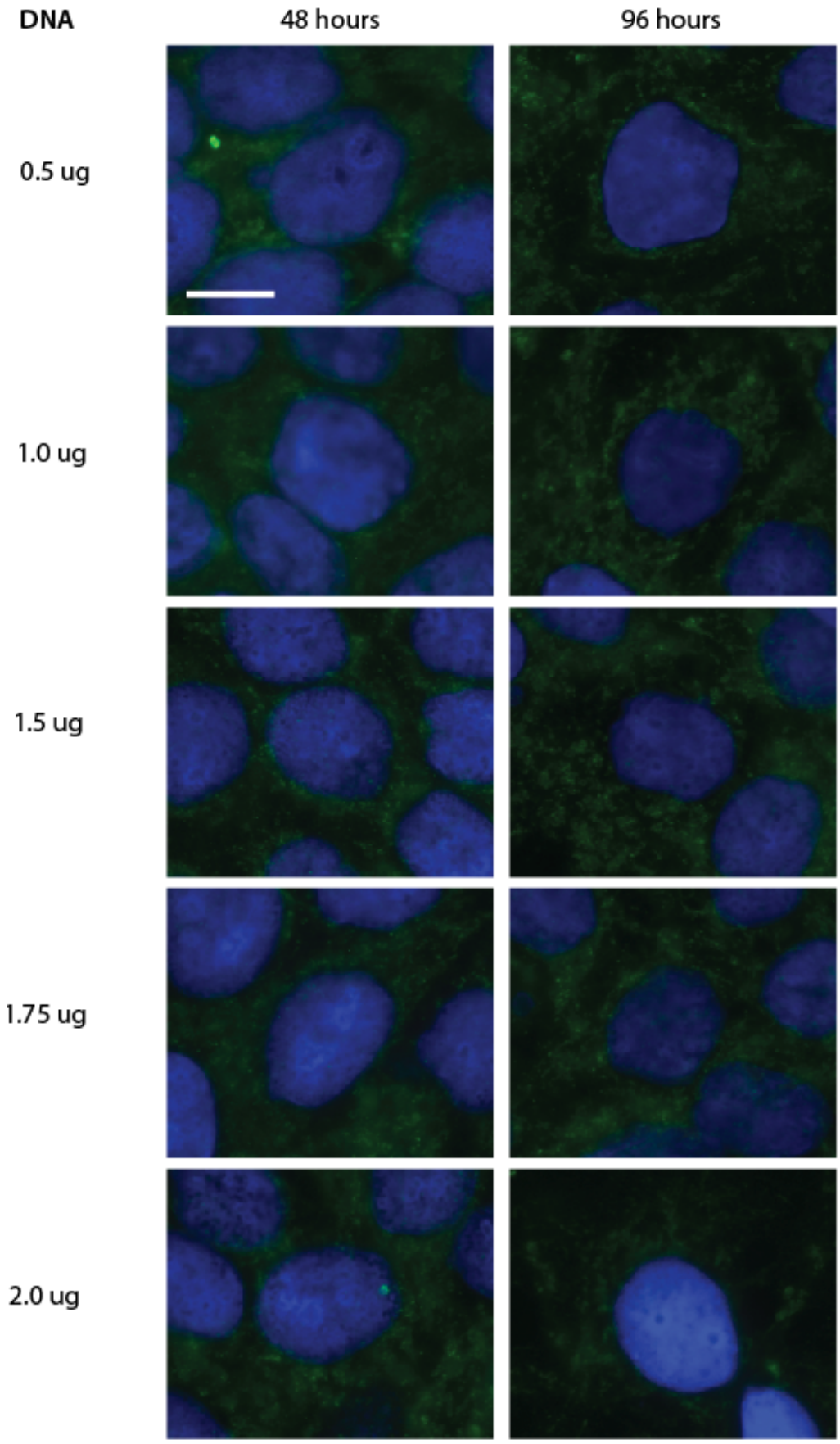



**Figure 3. 21 Immunofluorescent analysis of HaCaT cells transfected with myc-TMEM79 and incubated for 48 and 96 hours.** Cells were fixed and stained with an anti-myc mouse monoclonal antibody (green), showing no expression of myc-TMEM79. Nuclei were stained with DAPI (blue). Scale bar = 10 μm

Previously, transmission electron microscopy showed that TMEM79 protein was seen on the membranes of LG in terminally differentiating keratinocytes (Figure 3. 15). Markers for intracellular transport from the endoplasmic reticulum (ER) to Golgi (β-COPI), Golgi to ER (COPII) and the trans-Golgi network (TGN46) were used in immunofluorescent analysis, but no co-localization with TMEM79 was seen (Figure 3. 22).

Since the transfection efficacy using the myc-TMEM79 commercially available plasmid only yielded a detectable expression in <10% of cells, further investigation with an alternatively tagged TMEM79 was designed for subsequent experiments.

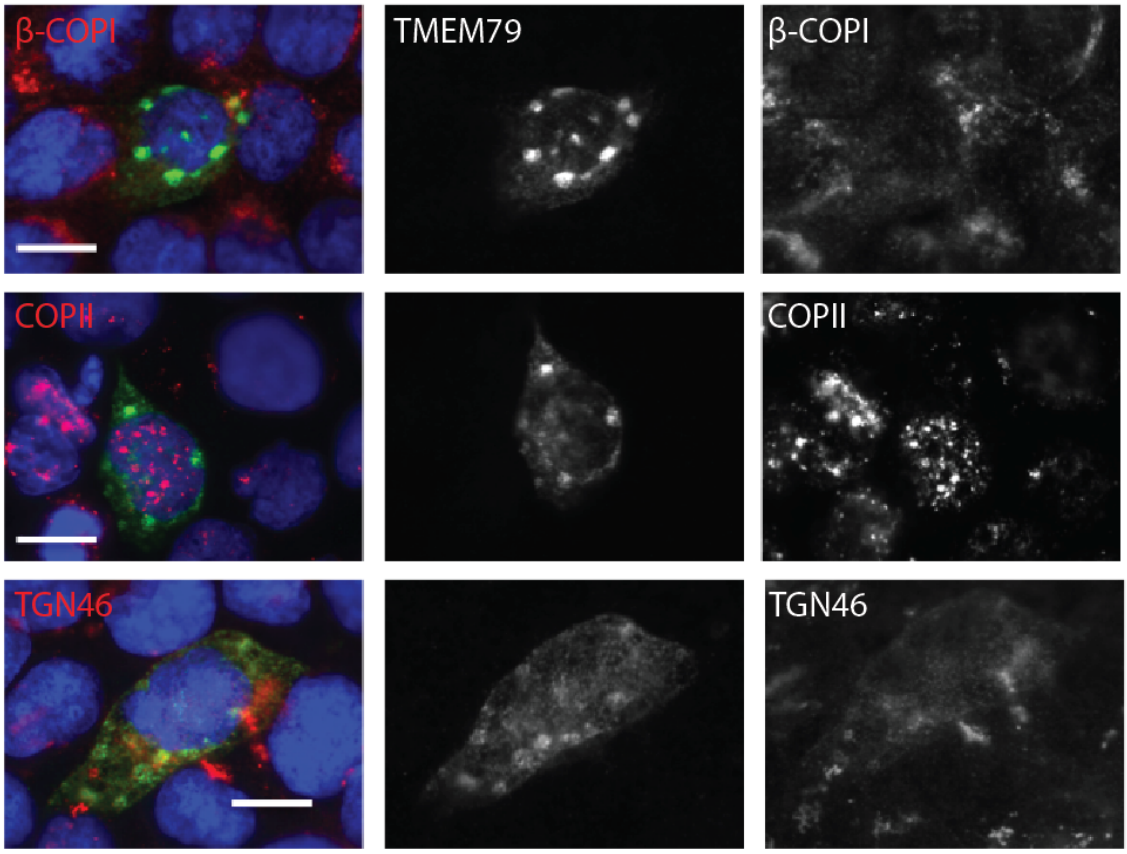



**Figure 3. 22 Co-immunofluorescent analysis of transfected HaCaT cells with markers for vesicle transport**. HaCaT cells were transfected with 0.5 µg of myc-TMEM79 and incubated for 24 hrs. No co-localization was seen with markers for vesicle transport from endoplasmic reticulum (ER; β-COPI) to Golgi, Golgi to ER (COPII) and the trans-Golgi network (TGN46). Leftmost column shows merged images for each co-immunofluorescent staining. Nuclei were stained with DAPI (blue). Scale bar = 10 µm

### 3.3.8.2. TMEM79/pEGFP overexpression

TMEM79 tagged with EGFP (using the vectors pEGFPC1 or pEGFPN1) gave a higher detectable rate of transfection (~60%) than the previously used myc-TMEM79 construct. In Figure 3. 24, HaCaT cells were transfected TMEM79/pEGFPC1 or TMEM79/pEGFPN1 and incubated for 24 hours, before fixing. Autofluorescence of EGFP expression was analyzed by fluorescence microscopy. TMEM79/pEGFPC1 produced small, vesicle-like structures at the lowest transfection concentration of 0.2 μg per well (in a 6 well dish), and the size of these structures increased when higher amounts of TMEM79/pEFGPC1 were transfected (Figure 3. 24; left panel). In contrast, transfection with TMEM79/pEGFPN1 gave a similar expression pattern as myc-TMEM79 transfected cells, showing a sprawling network with the occasional aggregate of protein (Figure 3. 24; right panel). Controls for 0.2 μg pEGFPC1 and pEGFPN1 vector plasmids and untransfected HaCaT cells were fixed and analysed by immunofluorescence microscopy, showing no specific EGFP expression in Figure 3. 23.

Total protein extracted from transfected cells was immunoblotted with anti-EGFP and anti-TMEM79 antibodies and gave a band of the expected size of ~64 kDa, confirming that the TMEM79-EGFP fusion constructs were indeed being expressed (Figure 3. 25). Control transfections using only the EGFP vector plasmid showed that each EGFP-expressed protein ran at a slightly different size; a band was detected at ~30 kDa in cells transfected with pEGFPC1 vector and a band of ~27 kDa from cells transfected with pEGFPN1 vector. This size difference is attributed to the size of the cloning nest sequence between the EGFP and TMEM79 sequence.

Co-immunofluorescent analysis with various markers was carried out for cells transfected with 0.2 μg of TMEM79-EGFP tagged fusion proteins to elucidate the origin/destination of these vesicles.

No co-localization with TMEM79 was seen with β-COPI and COPII. Both are markers for vesicular retrograde/anterograde transport between the ER to the Golgi complex respectively. β-actin staining for both tagged fusion proteins showed a normal cytoskeletal actin network. TGN46 is a marker for the trans-golgi network, and showed no co-localization with vesicles from transfection of TMEM79/pEGFPC1. LAMP2 (lysosomal-associated membrane protein 2) staining shows no co-localization of TMEM79 indicating that it is not associated with lysosomes. EEA1 (early endosome antigen 1) and Rab 5 (marker for early endosomes) both show no association of TMEM79 with either early endosomes, or vesicles associated with endocytosis. LC3, a marker for autophagosomes, show partial co-localization with the TMEM79 aggregates seen in cells transfected with TMEM79/pEGFPC1 (Figure 3. 26).

The network-like staining pattern of TMEM79/pEGFPN1 protein did not show any co-localization with any of the markers used (Figure 3. 27).

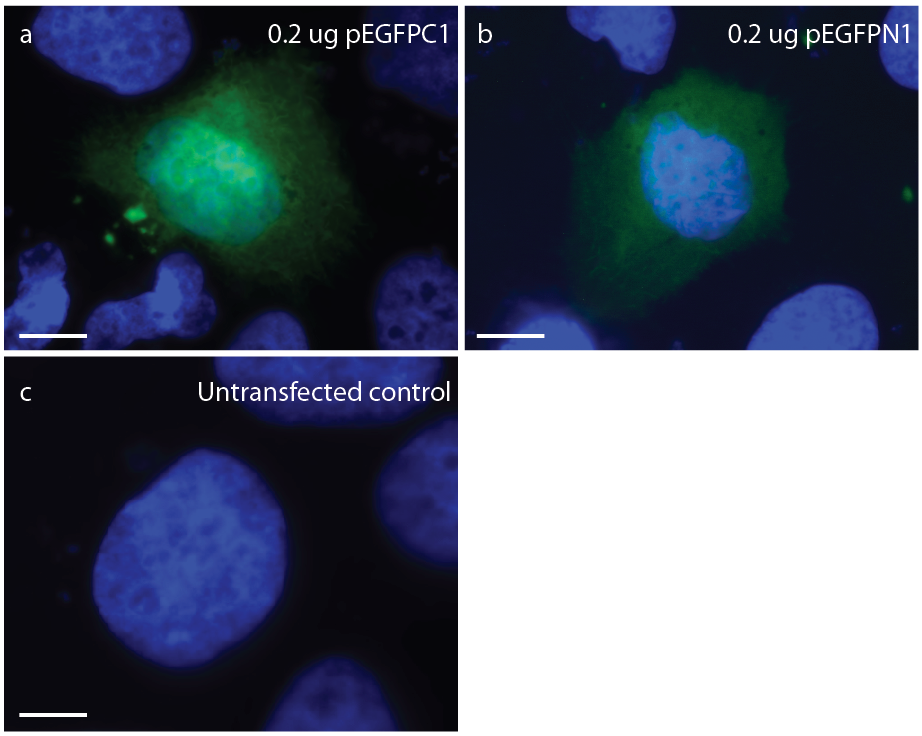



**Figure 3. 23 Autofluorescent analysis of HaCaT cells transfected with EGFP1 vectors.** HaCaT cells were transfected with 0.2 µg of pEGFPC1/pEGFPN1 vector plasmids and incubated for 24 hours showing no specific expression of EGFP (a and b; green). Untransfected cells showed no EGFP expression (c). Nuclei were stained with DAPI (blue). Scale bar = 10 µm

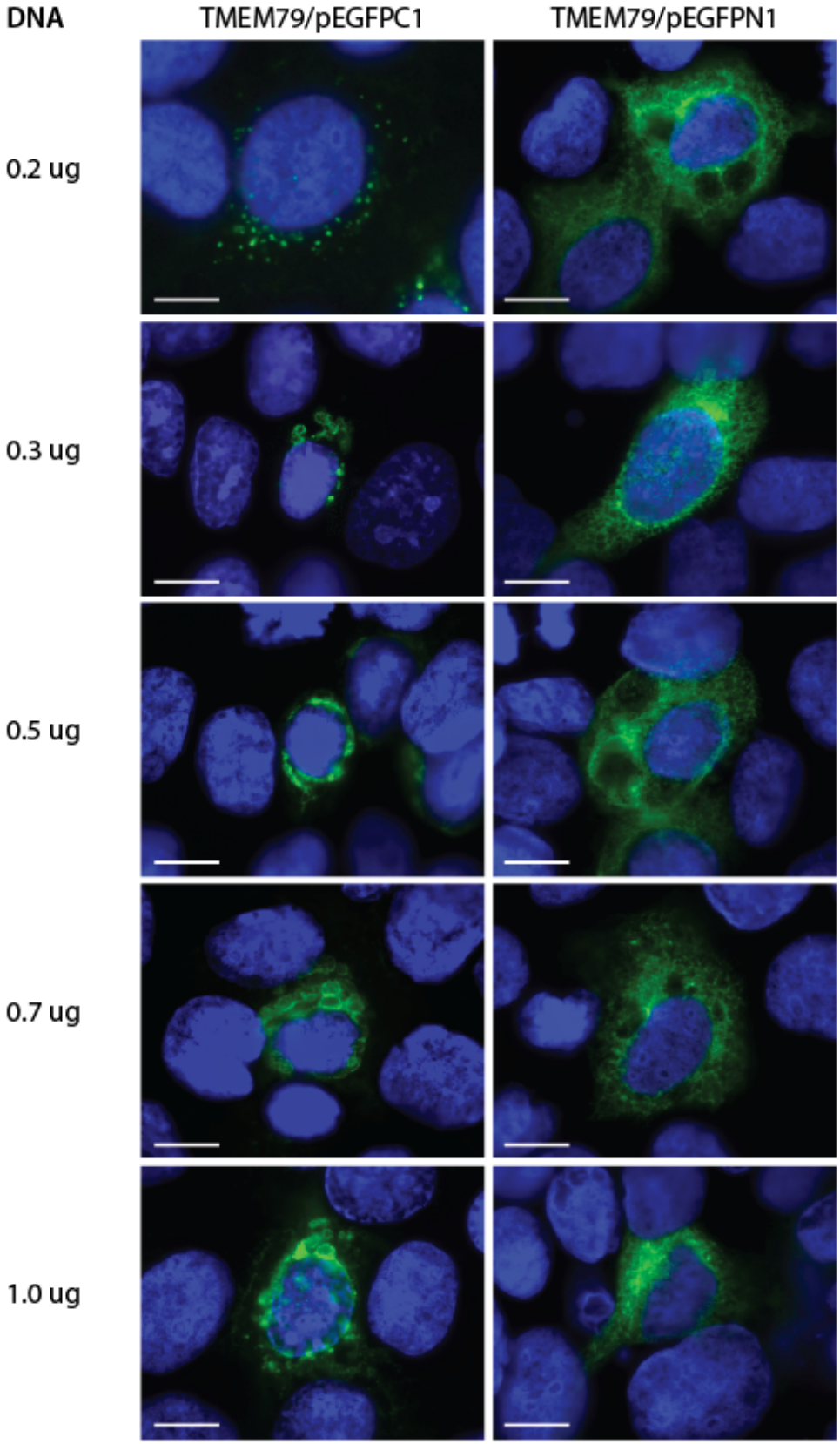



**Figure 3. 24 Autofluorescent analysis of HaCaT cells transfected with TMEM79/pEGFPC1 or TMEM79/pEGFPN1.** Cells were incubated for 24 hrs. Transfection with TMEM79/pEGFPC1 resulted in small cytoplasmic aggregates, which increased in size with transfected DNA concentration (green; left panel). TMEM79/pEGFPN1 transfected HaCat cells displayed a network-like expression pattern (green; right panel). Nuclei were stained with DAPI (blue). Scale bar = 10 μm

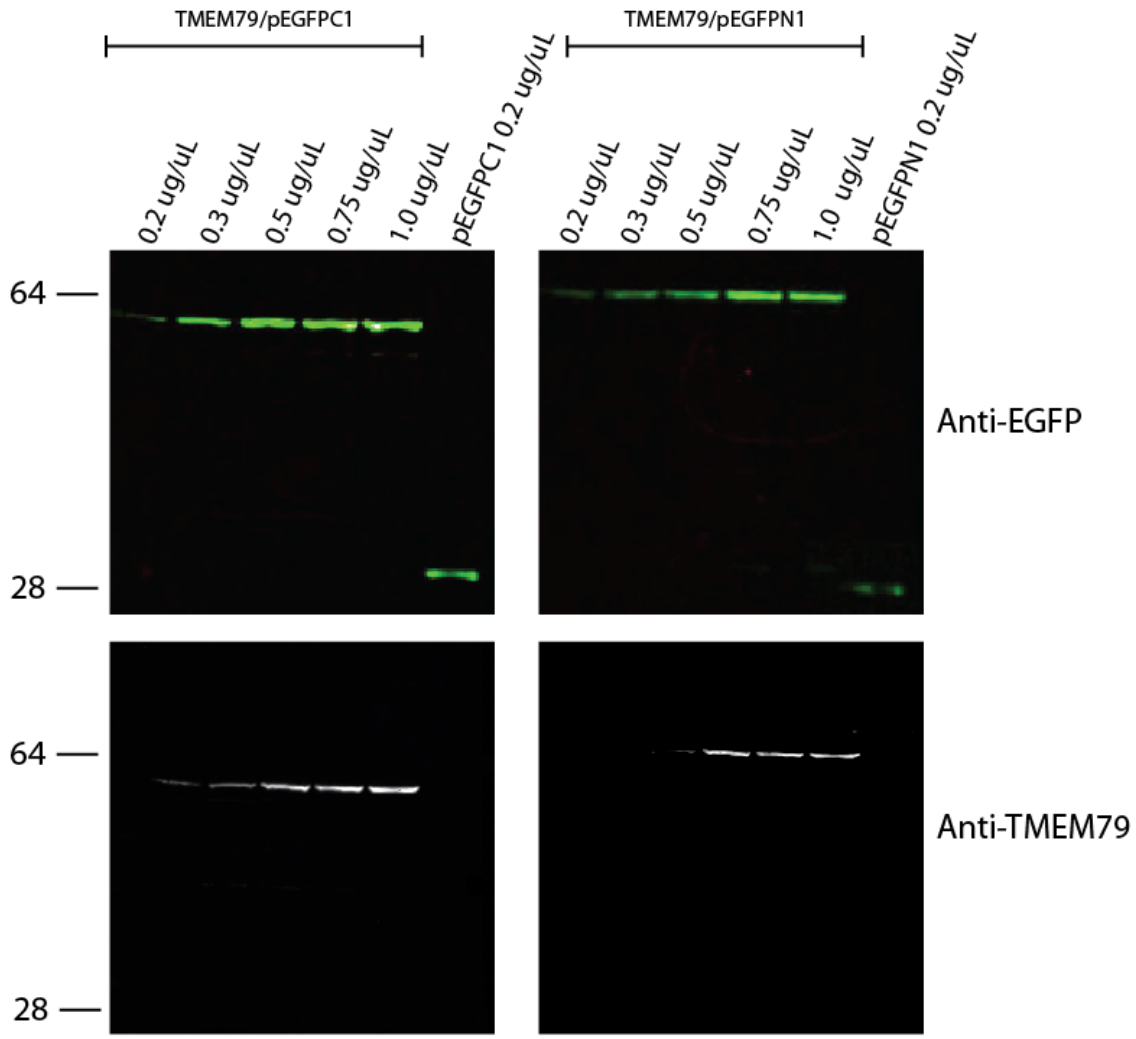



**Figure 3. 25 Immunoblots of total protein extracts from transfected HaCaT cells.** Cells transfected with TMEM79/pEGFPC1 (left blots) and TMEM79 pEGFPN1 (right blots) were incubated for 24 hours. Anti-TMEM79 rabbit polyclonal antibody was used to detect TMEM79 and anti-EGFP mouse monoclonal antibody was used to detect EGFP. A band was detected at the expected protein size ~64 kDa. Transfection with EGFP vector plasmids as controls shows a band ~30 kDa for EGF.

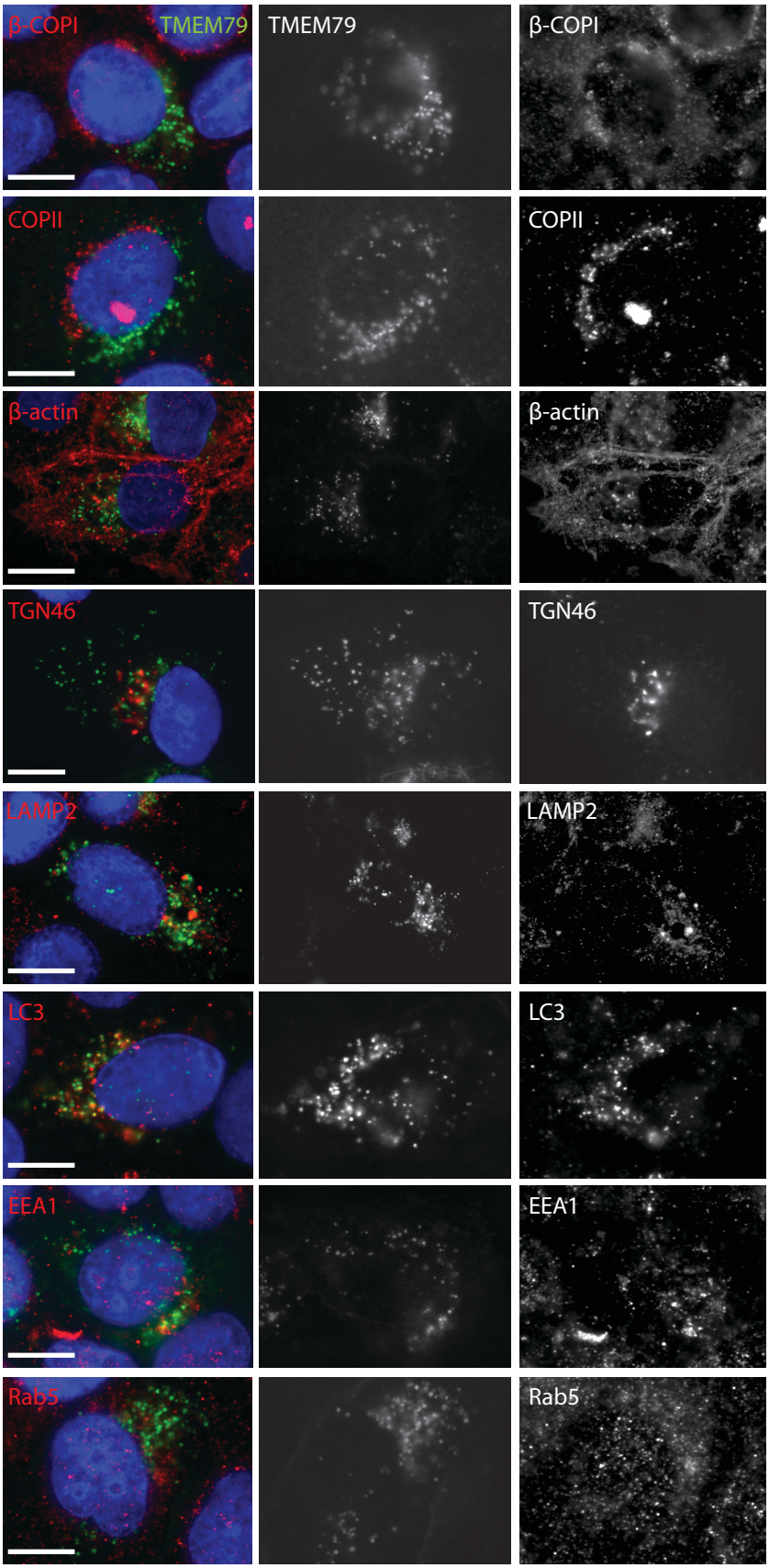



**Figure 3. 26 Co-immunofluorescent analysis of HaCaT cells transfected with TMEM79/pEGFPC1 together with markers for Golgi transport and autophagy.** Cells transfected with 0.2 µg TMEM79/pEGFPC1 were incubated for 24 hrs. Leftmost column show merged images for each co-immunofluorescent staining. Partial co-localization of TMEM79 was seen with LC3, a marker for autophagosomes. Nuclei were stained with DAPI (blue). Scale bar = 10µm

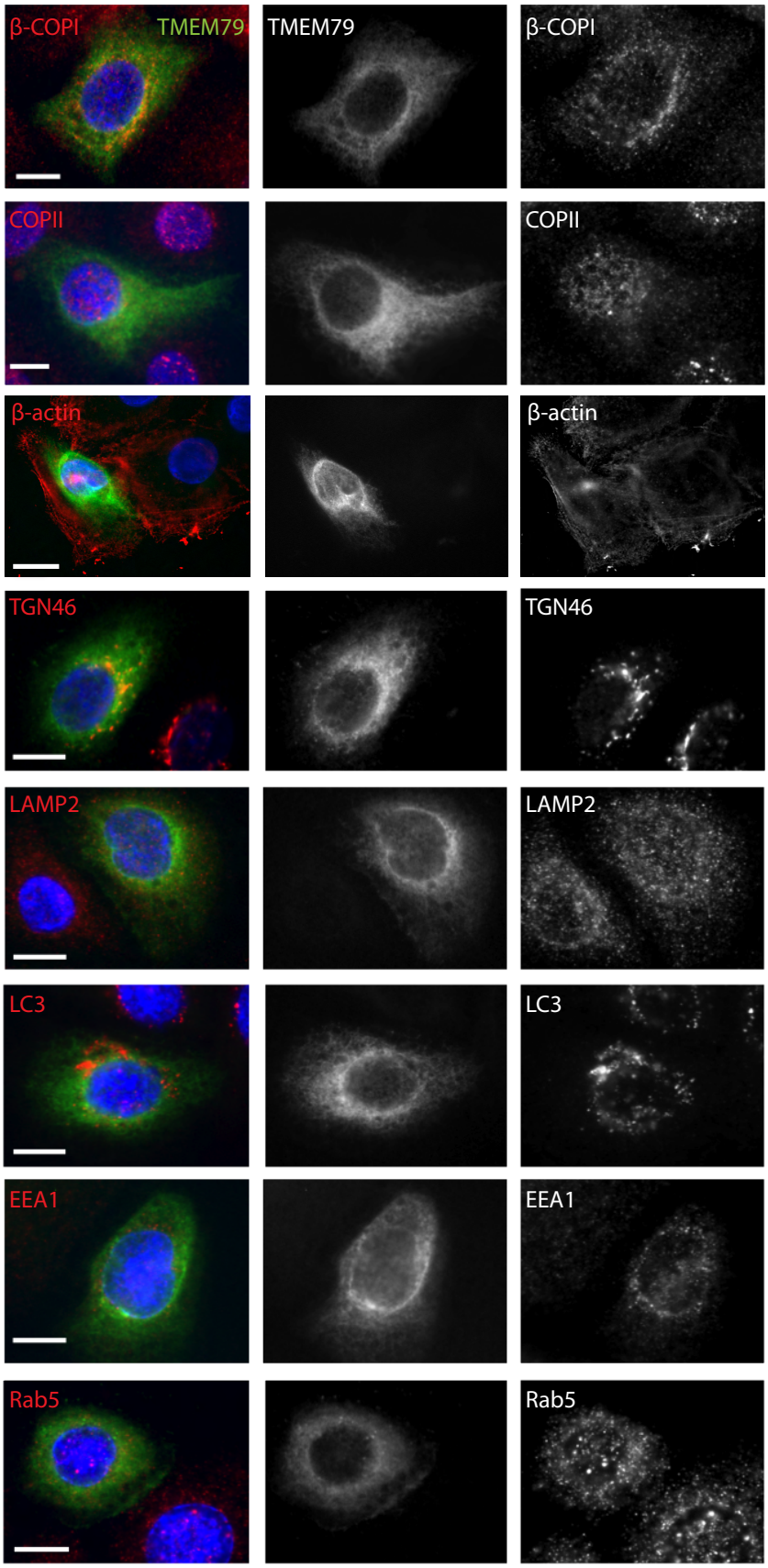



**Figure 3. 27 Co-immunofluorescent analysis of HaCaT cells transfected with TMEM79/pEGFPN1 together with markers for Golgi transport and autophagy.** Cells transfected with 0.2 µg TMEM79/pEGFPC1 were incubated for 24 hrs. Leftmost column show merged images for each co-immunofluorescent staining. No co-localization was observed with TMEM79. Nuclei were stained with DAPI (blue). Scale bar = 10µm

#### 3.3.8.3. Transmission electron microscopy of TMEM79/pEGFPC1 transfected HaCat cells

TMEM79/pEGFPC1 transfected HaCaT cells were selected for further analysis using transmission electron microscopy due to partial co-localization of aggregates with autophagosome marker LC3 (seen in Figure 3. 26). Transiently transfected HaCaT cells with 0.2 μg of TMEM79/pEGFPC1 were visually identified on gridded coverslips by fluorescence microscopy and flagged up for transmission electron microscopy analysis. Autophagosomes are double-membrane bound structures that engulf proteins from the cytoplasm that are targeted for degradation by fusing with lysosomes to form an autophagolysosomes. Distinct double-membrane bound structures could be seen clearly in clusters on one side of the nucleus at lower magnification (Figure 3. 28a, c, e), and the black arrowheads indicate the region that is shown at a higher magnification in the panel on the right (Figure 3. 28b, d, f). Early autophagosomes are seen as the crescent-shaped structures in Figure 3. 28b and d (black arrows), and mature autophagosomes can clearly be seen in Figure 3. 28d and f (white arrows), where the double bilayer membrane has fused together to engulf the material meant for degradation. The sizes of these autophagosomes correspond with the aggregates of transfected protein seen by immunofluorescence, suggesting that the overexpression of TMEM79 is leads to the rapid degradation of the protein, and a more stable expression system is needed for future experiments.

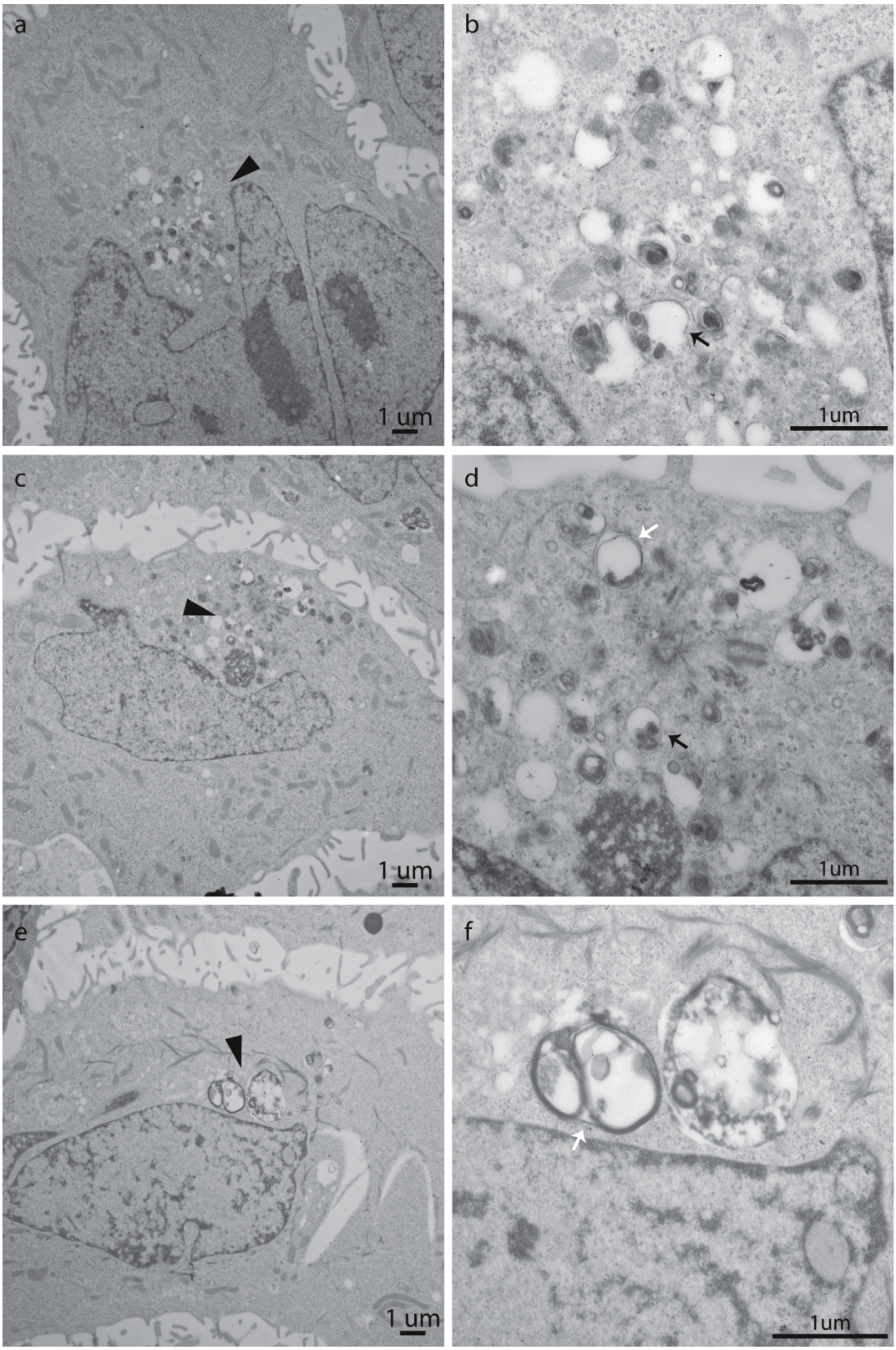



**Figure 3. 28 Transmission electron microscopy of HaCaT cells transfected with TMEM79/pEGFPC1.** Cells were incubated for 24 hours and a large number of autophagosomes were seen in transfected cells. The regions indicated by black arrowheads (a, c and e) are magnified in the right panel (b, d and f); young autophagosomes (b and d; black arrows) are easily distinguished from mature autophagosomes (d and f; white arrows). Over-expressed TMEM79 constructs appear to be degraded via autophagocytosis.

#### 3.3.8.4. Site-directed mutagenesis and overexpression of p.Y283*/pEGFPC1

The human equivalent of the matted mutation, p.Y283*, was introduced by site-directed mutagenesis into the *TMEM79* coding sequence by changing a single base, T to A at cDNA position 847, in order to introduce a premature stop codon. This mutation occurs within the third transmembrane domain, and the mutant protein lacks its C-terminus. To test whether the matted mutant construct would give a different expression pattern, transfection of this mutant construct in HaCaT cells were performed, and showed two different patterns of expression (Figure 3. 29.) Cells either displayed a honeycomb-like structure that increased in size with the amount of transfected construct (Figure 3. 29, Pattern 1) or small aggregates of protein (Figure 3. 29, Pattern 2) that were similar to the aggregates seen in HaCaT cells transfected with full-length TMEM79/pGEFPC1 (Figure 3. 24; left panel, page 118). An equal percentage of transfected cells were seen displaying each pattern. Controls for HaCaT cells transfected with 0.2 µg pEGFPC1 vector plasmid and untransfected cells showed no EGFP expression pattern (Figure 3. 23, page 117). Analysis of the co-localization of the p.Y283*/pEGFPC1 construct with the same markers for Golgi transport and autophagy as performed earlier for the wildtype TMEM79/pEGFPC1 showed no co-localization with any of the markers, apart from LC3 (Figure 3.28). This pattern of co-localization with LC3 suggests that overexpression of the mutant TMEM79 may also result in its degradation by autophagocytosis.

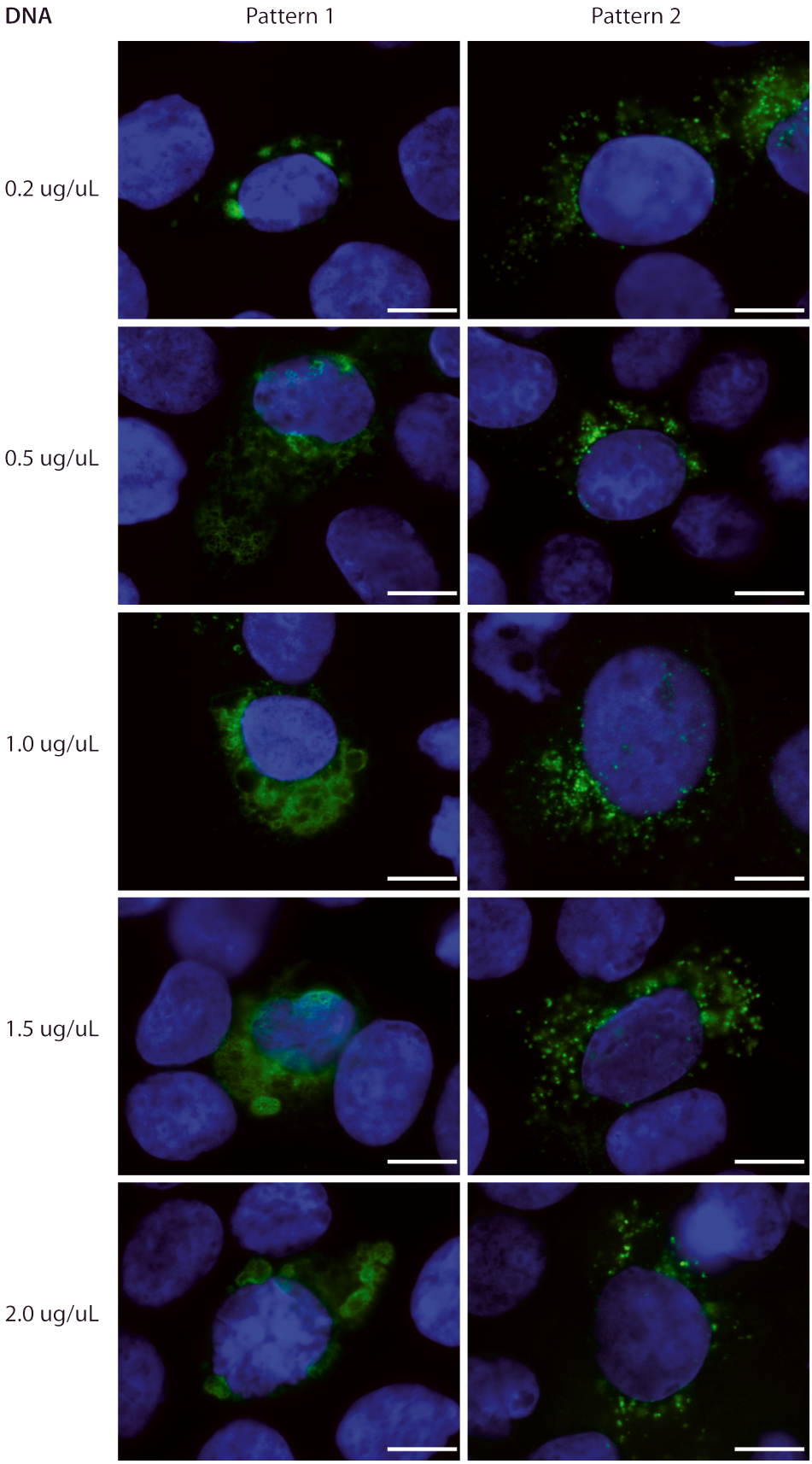



**Figure 3. 29 Autofluorescent analysis of HaCaT cells transfected with p.Y283*/pEGFC1.** Cells were transfected with 0.2 µg of p.Y283*/pEGFPC1 and incubated for 24 hours. Two different expression patterns were seen (green). Large vesicle-like structures seen in the left panel, and small aggregates seen in the right panel. Nuclei were stained with DAPI (blue). Scale bar = 10 µm

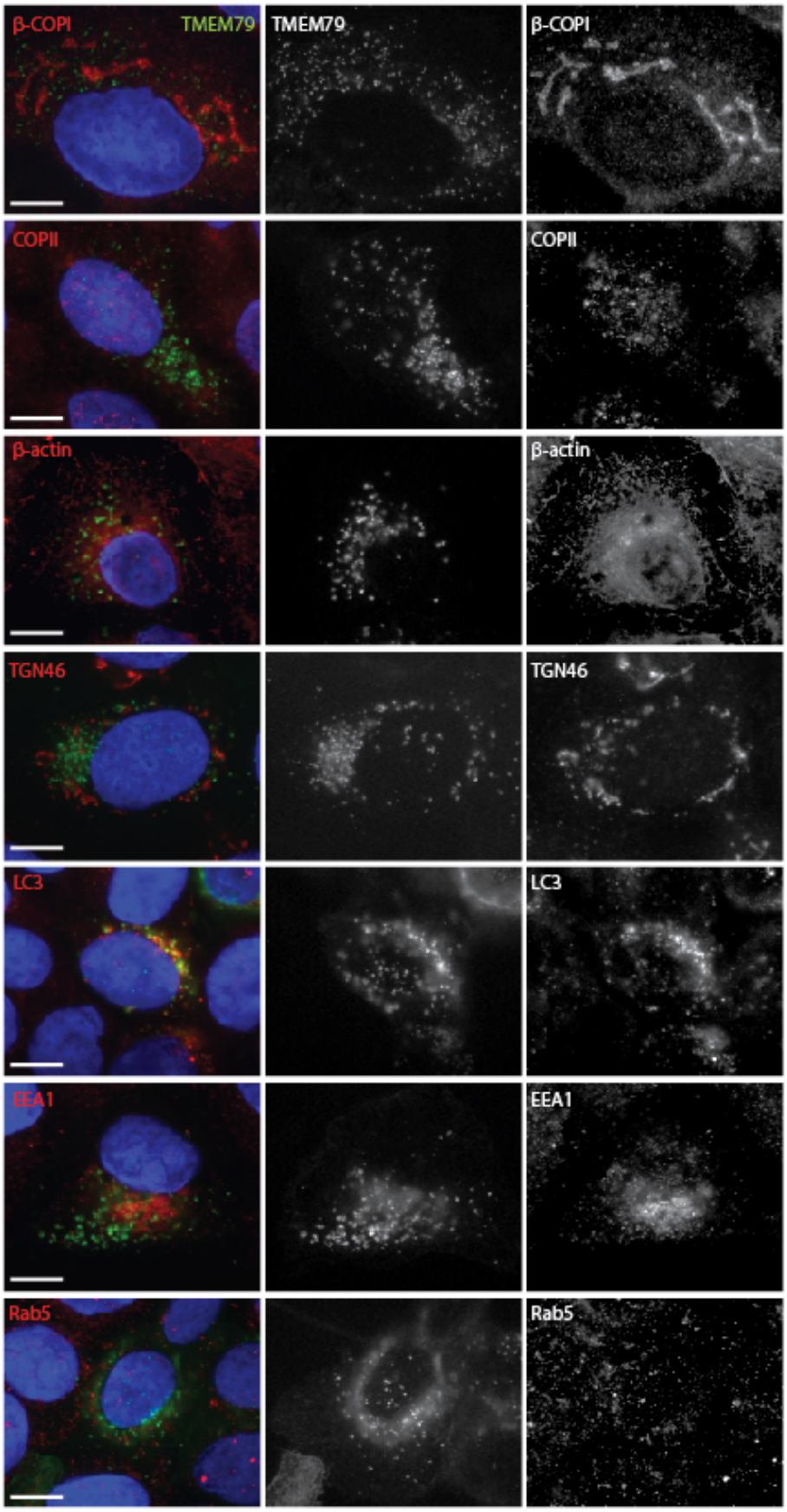



**Figure 3. 30 Co-immunofluorescent analysis of HaCaT cells transfected with p.Y283*/pEGFPC1 and stained with markers for Golgi transport and autophagy.** Partial co-localization was seen with LC3, a marker for autophagosomes. Nuclei were stained with DAPI (blue). Scale bar = 10μm

### 3.3.8.5. Co-overexpression of TMEM79/pEGFPC1 and p.Y283*/pEGFPC1

0.2 µg of TMEM79/pEGFPC1 and 0.2 µg of p.Y283*/pEGFPC1 were transfected together in the same well to test for a dominant negative effect. It was hypothesized that the presence of p.Y283*/pEGFPC1 may antagonize the wildtype TMEM79 protein, leading to an altered expression pattern. Cells transfected with both plasmids displayed a similar honeycomb-like structure (Figure 3. 31) to cells transfected with only the mutant p.Y283*/pEGFPC1 construct (Figure 3. 29; Pattern 1). Based on the expression pattern seen below, it is not possible to determine whether the mutant p.Y283*/pEGFPC1 has a dominant negative effect.

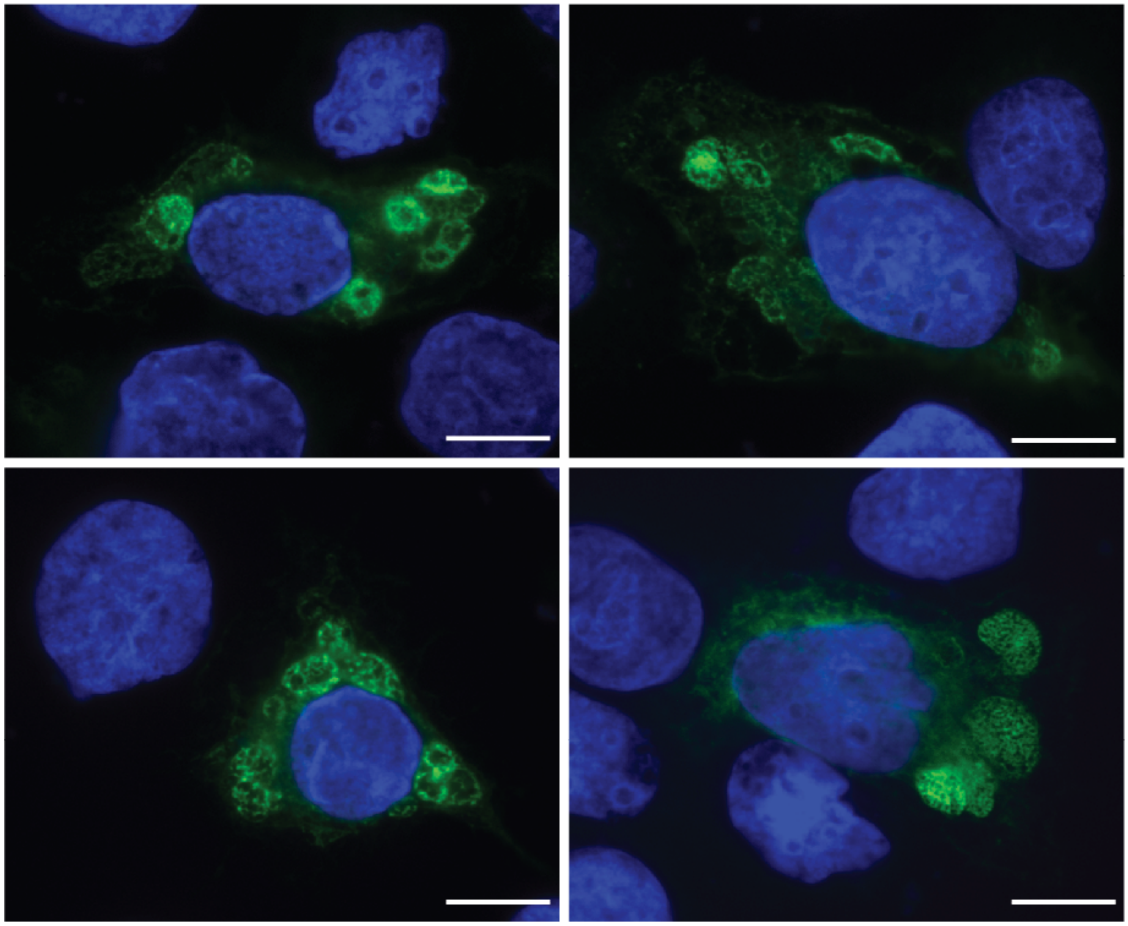


**Figure 3. 31 Autofluorescent analysis of HaCaT cells co-transfected with TMEM79/pEGFPC1 and p.Y283*/pEHGPC1.** Cells were transfected with 0.2 µg of TMEM79/pEGFPC1 and 0.2 µg of p.Y283*/pEGFPC1 and incubated for 24 hours. Honeycomb-like structures were seen in transfected cells (green). Nuclei were stained with DAPI (blue). Scale bar = 10 µm

### 3.3.9. Transient overexpression of TMEM79/pEGFP in human primary keratinocytes (HPK)

As the previous experiments were performed on an immortalized keratinocyte cell line (HaCaT cells), HPK were used next to see if there was any difference in expression. HPK transfected with TMEM79/pGEFPN1 and incubated for 24 hours showed a network-like staining pattern (Figure 3. 34) when transfected with 0.2 µg of plasmid per well (in a 6 well dish), similar to HaCaT cells transfected with full-length TMEM79/pEGFPN1 (Figure 3. 24; right column, page 118), but this expression pattern gradually became fainter as the concentration of transfected plasmid per well increased and were replaced by large aggregates. Visually, a small percentage of cells were seen to have this network-like staining, but the majority of the cells showed large protein aggregates, which varied from 1 – 5 µm in diameter and this phenotype was seen in HPK transfected with both TMEM79/pEGFPC1 (Figure 3. 33) and TMEM79/pEGFPN1 (Figure 3. 34), when increasing concentrations of protein constructs were transfected into HPKs. Controls transfected with pEGFPC1 and pEGFPN1 show no specific expression pattern of EGFP, and untransfected controls were completely negative for any EGFP expression (Figure 3. 32).

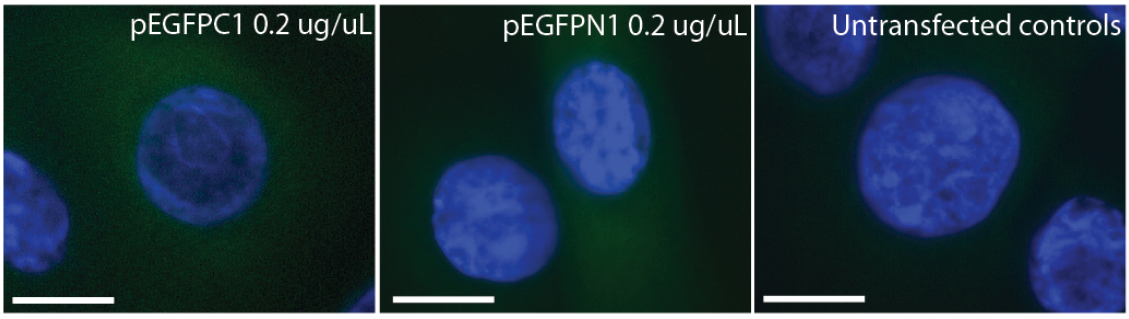



**Figure 3. 32 Human primary keratinocytes transfected with pEGFPC1 or pEGFPN1 vector.** HPKs were transfected with 0.2 µg of pEGFPC1 or pEGFPN1 vector for 24 hours show no specific autofluorescent EGFP expression. Untransfected controls did not show any EGFP expression. Nuclei were stained with DAPI (blue). Scale bar = 10 µm

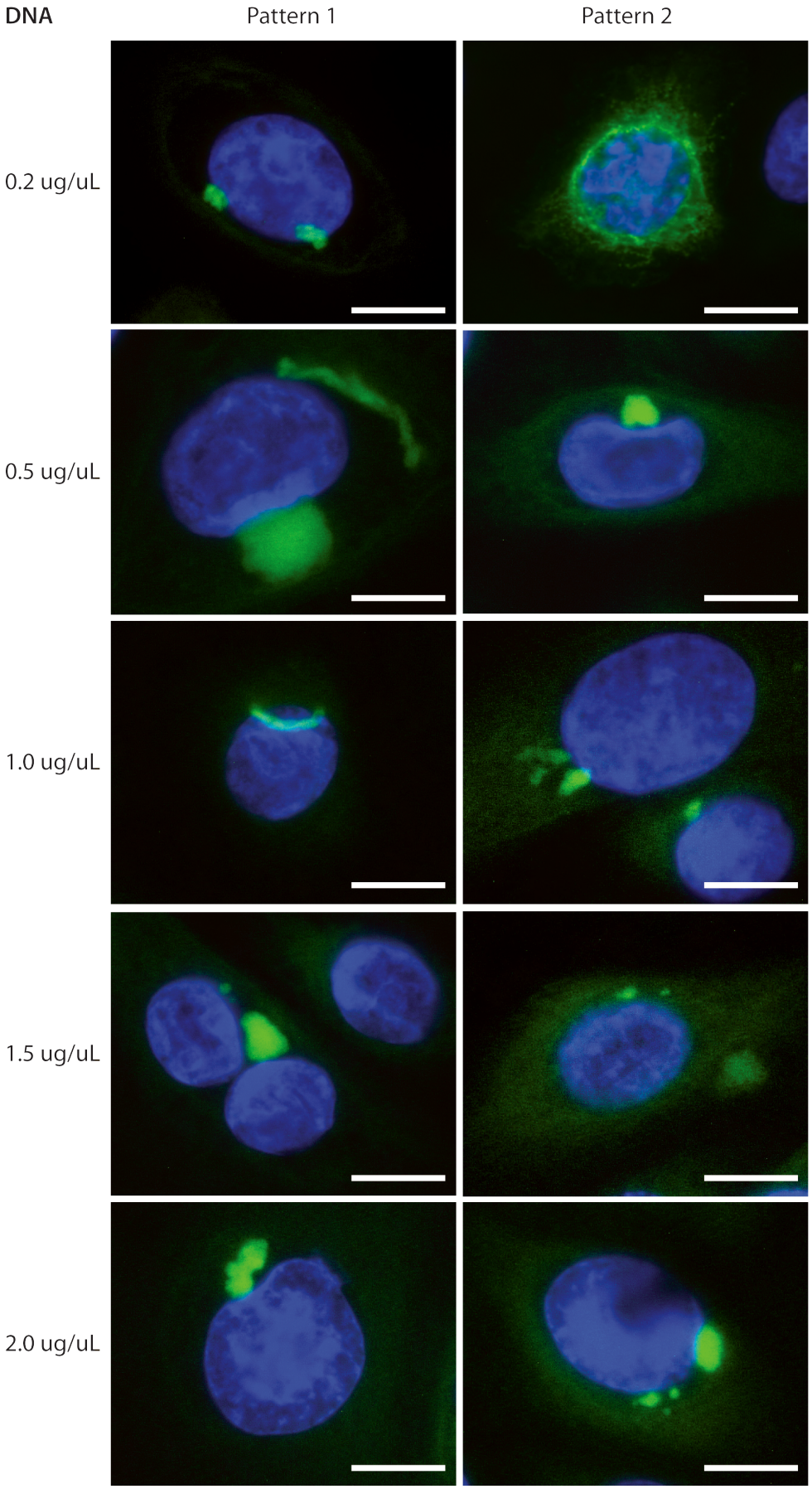



**Figure 3. 33 Autofluorescent analysis of human primary keratinocytes transfected with TMEM79/pEGFPC1.** Cells were incubated for 24 hours (green) and showed large aggregates. Nuclei were stained with DAPI (blue). Scale bar = 10 μm

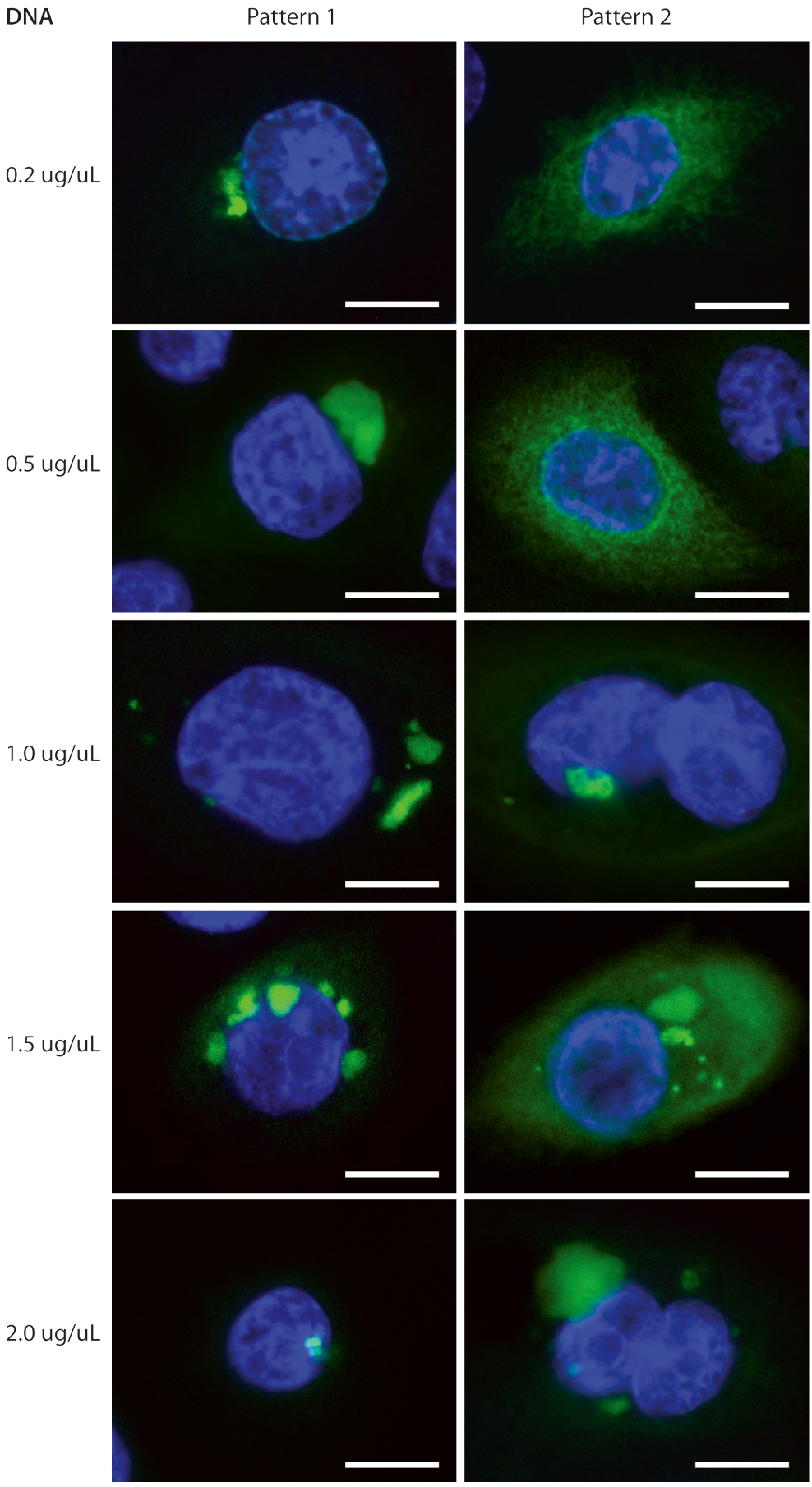



**Figure 3. 34 Autofluorescent analysis of human primary keratinocytes transfected with TMEM79/pEGFPN1.** Cells were incubated for 24 hours (green). Nuclei were stained with DAPI (blue). Scale bar = 10 μm

## 3.4. DISCUSSION

The location of *TMEM79* within the EDC on human chromosome 3 suggests a high likelihood that this protein may be involved with epidermal differentiation and maturation. Immunofluorescence of human abdominal skin samples showed that TMEM79 was expressed in the upper granular layers and partially co-localized with the upper suprabasal marker keratin 10, and desmosomal marker desmoplakin 1/2. Known markers for lamellar granules (LG) such as corneodesmosin and kallikrein 8 were seen in the same upper granular layers as TMEM79. Further, immunofluorescent analysis of hair follicles localized TMEM79 to the Huxley layer of the IRS, and TMEM79 could also be associated with the LG structures of the IRS. Future experiments involving immune-gold labeling TMEM79 in hair follicle sections should be performed to confirm the subcellular location of TMEM79 within the cells of the IRS. In human epithelia, expression studies of multiple tissues show that TMEM79 transcript variant 1b is seen to have the highest expression in stratified tissues such as prostrate, cervix, skin and tongue. This suggests a possibility that TMEM79 transcript variant 1b may be involved in the differentiation process that is unique to stratified tissues. Transmission electron microscopy shows evidence for possible involvement with terminal differentiation of the epidermis, where TMEM79 is associated with the membranes of LG at the SG-SC junction. These LG are seen as membrane-bound vesicular or tubular structures within keratinocytes of the upper granular layer, located close to the plasma membrane and are thought to contain distinct different cargoes, including glucosylceramides, kallikrein-related peptidases 5, 7 and 8, cathepsin D and corneodesmosin. Although the mechanisms for transport of various LG cargos are still unknown, it has been suggested that segregation of these contents would allow control in concentration and secretion of these

different cargoes (Dannies, 2001). CDSN has been documented to be expressed slightly earlier than KLK8 in keratinocytes undergoing terminal differentiation (Ishida-Yamamoto et al., 2004), which agreed with the images of gold-labeled ultra-thin sections of human epidermis shown in this chapter, with KLK8 being located in LGs further away from the apical face at the SG-SC interface whilst CDSN was already located at the desmoglea of desmosomes. A potential function for TMEM79 could be as a label for a specific type of LG, as both CDSN and KLK8 have been shown to be located within LG that are not labeled with TMEM79. This form of labeling could enable LG cargoes to be sorted and released at different times during differentiation.

Multiple difficulties were encountered while trying to detect endogenous TMEM79 from human epidermis by Western blotting. This could be attributed to the complications of working with a 5-pass transmembrane protein, which makes it more difficult to extract compared to a cytosolic protein. Samples also could not be sonicated, as this caused membranous proteins to go into insoluble aggregates. Lowering the denaturation temperature to 37°C and incubating for an hour gave the best results. The antibody used to detect TMEM79 was the same one used for immunofluorescence and TEM, and it is possible that the antibody does not work as well for Western blotting, as for the other techniques. Epitope conformation could also be different on a blot, than in tissue, and this could affect antibody binding.

Overexpression of TMEM79 in cell monolayers showed different localization patterns, and this could have been due to the different tags used and whether they were N-terminal or C-terminal. The expression patterns for TMEM79/pEGFPN1 and myc-TMEM79 are consistently similar and it could deduced that the localization of TMEM79 is affected by the position of tags on the C-terminus of the

TMEM79 construct, by blocking any potential protein-binding or recognition sites. At the 48 and 96 hour time points after transfection, no expression of TMEM79 constructs were seen in HaCaT cells and further investigation at the 24 hour time point showed some indication that the overexpression of TMEM79 constructs caused the aggregation of these proteins that may have resulted in degradation via autophagy. Another possibility might be that the cells that are transfected die if the protein degradation machinery of the cell becomes overloaded.

Co-immunofluorescence suggests that a fraction of the aggregates co-localize with LC3 and this partial co-localization could be due to the different maturation stages of an autophagosome, as there might be obstruction of the LC3 epitope after an autophagosome fuses with a lysosome to form an autophagolysosome. To confirm that the double membrane structures are indeed autophagosomes, immuno-gold labeling of transfected cells with antibody against LC3 should be carried out.

Even though *TMEM79* message was seen in both HaCaT cells and HPKs, no endogenous TMEM79 was seen in either cell line by immunofluorescence. Transiently transfected TMEM79 protein aggregated in large clumps for both cell lines, indicating that TMEM79 protein is not normally expressed within these cells, and overexpression is causing TMEM79 protein aggregation leading to eventual degradation. This result correlates well with the expression of TMEM79 in normal human skin. TMEM79 is found only in the upper granular layers, where the process of terminal differentiation would have already been well initiated, and so would not be normally expressed in a cultured monolayer. For future dominant negative experiments, transfections with differently tagged mutant protein constructs and wildtype protein constructs should be used so that positively transfected cells with both plasmids can be identified by co-immunofluorescent staining.

The DMEM media used to culture HaCaT cells contained a high concentration of 1.8mM calcium chloride, which is important for keratinocyte stratification. For all experiments performed in this chapter, cells were grown to 80% - 90% confluency, and it might be more useful to leave the HaCaT cells to stratify before experimenting with TMEM79 overexpression, since TMEM79 is only expressed at the SG-SC interface. HPKs were cultured in a $Ca^{2+}$ free CnT media, which would not have encouraged stratification. This might explain why overexpression of TMEM79 constructs were consistently being autophagocytosed.

In conclusion, this chapter has confirmed that TMEM79 is expressed in keratinocytes at the SC-SG interface and in cells of the IRS of the hair follicle, with specific localization to the membranes of lamellar granules. This protein appears to have high identity matches across species, suggesting that these conserved residues may have an important functional role. Analysis of possible transcript variants of human *TMEM79* identified *TMEM79* transcript variant 1b (uc009wrw.3) as the most abundant transcript variant in the skin, prostrate, tongue and mammary gland. *TMEM79* transcript variant 1a (uc010phi.2) was found to be expressed in various tissues using a multiple cDNA tissue array. Using bioinformatics analysis, TMEM79 was predicted to have 5 transmembrane regions, and these regions were conserved across species such as naked mole rat, chicken and zebrafish. Co-localization analysis of transfected HaCaT cells and HPK showed no specific expression pattern with myc-tagged and EGFP-tagged TMEM79 protein. Overexpression of TMEM79-fusion constructs in a monolayer culture results in protein degradation via autophagy. Stratification of cells either by inducing with high calcium or a more stable expression assay system needs to be developed, where cells would be able to express TMEM79 at low levels in a non-transient manner.

# CHAPTER 4

# Association of *TMEM79* with human disease

## 4.1. AIMS OF CHAPTER

The overt allergic phenotype of the *Matt*$^{ma/ma}$ mouse suggests a role for TMEM79 in skin barrier function. Naturally, the next step is to explore the landscape of TMEM79 within various atopic human populations and to identify if there is any association with allergic disease. In the previous chapter, ultrastructural investigations into the TMEM79 protein revealed that it is located within the membranes of lamellar granules, and could result in abnormal LG secretion/transport when normal function is disrupted during disease states. To date, several ichthyosiform disorders have been linked to abnormal lipid lamellar granule secretion, and these have been linked to transmembrane transporter proteins, such as ABCA12 in Harlequin ichthyosis, VPS33B in ARC syndrome and lipoxygenases in NCIE.

In this final chapter of results, several different human populations were screened to uncover SNPs within the novel *TMEM79* gene, and whether these may have any potential impact on protein function in contributing to atopic susceptibility. The aim of this chapter was to screen pre-existing populations of patients who share clinical phenotype with the *Matt*$^{ma/ma}$ mouse without any known mutations in other genes, and to investigate for TMEM79 variant impact.

## 4.2. MATERIALS AND METHODS

### 4.2.1. Patient cohorts for *TMEM79* sequencing

Human genomic samples were obtained with informed written consent from patients, together with approval from the relevant local authorities that complied with the Declaration of Helsinki Principles.

A total of 9 patients with Netherton Syndrome-like symptoms such as ichthyosiform erythroderma and trichorrhexis invaginata (Figure 4. 1) were recruited and screened by Professor Eli Sprecher (The Tel Aviv Sourasky Medical Center, Israel) (Table 4. 1), but were found to be negative for any SPINK5 mutations.

Atopic dermatitis (AD) Irish patients were diagnosed and recruited under the supervision of Dr. Alan Irvine and Dr. Rosemarie Watson (Our Lady's Children's Hospital for Sick Children, Ireland) (Table 4. 2). A second group of Irish atopic dermatitis patients with IgE > 10,000 were recruited from Ireland by Dr Jonathan Hourihane (Our Lady's Children's Hospital for Sick Children, Crumlin, Ireland) and Dr Lesley Anne Murphy (University Galway Hospital, Ireland). Population controls were obtained from the Irish Blood Transfusion Service Biobank.

A cohort of 55 patients with inflammatory eye disease that were diagnosed with blepharokeratoconjunctivitis (BKC; Table 4. 2) was recruited by Professor Johnny Moore (Cathedral Eye Clinic, Northern Ireland) and Dr. Tara Moore (University of Coleraine, Northern Ireland). All patients were diagnosed with an early onset of blepharitis (indicating genetic predisposition) and anterior and/or posterior conjunctival inflammation.

90 AD patients and 82 ethnically matched population controls from the Xhosa population from South Africa were diagnosed and recruited by Dr. Fatema Thawer-Esmail (The University of Cape Town, South Africa) (Table 4. 2).

| Patient | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Gender | M | M | F | F | Family of patient 4 | | M | Family of patient 7 | M |
| Age (years) | Died age 8 | 4 | 32 | 29 | | | 40 | | 26 |
| Eczema severity score (1-3)* | 3 | Unknown | 1 | 3 | | | 1 | | 2 |
| Ethnic origin/nationality | Mexican | Australian-Egyptian | Israel/Ashkenazi/Jew | Jordanian | | | Georgian | | Arab/Muslim/Israel |
| IgE levels (kU/L)**) | Unknown | Unknown | ~300 | Unknown | | | Around 1000 | | Normal |
| Erythroderma | Yes | Yes | No | Yes | | | No | | Yes |
| Occular inflammation | Yes | Yes | No | Yes | | | No | | Yes |
| Hair anomaly | Hypotrichosis | Brittle hair | Hypotrichosis | Pili torti and trichorrhexis invaginata | | | Trichorrhexis invaginata | | None |
| Metabolic wasting | Yes | Yes | No | No | | | No | | None |
| Asthma | No | Yes | Yes | No | | | Yes | | No |

**Table 4. 1 Clinical and phenotypic details of the 9 cases with Netherton Syndrome-like features.**

* Eczema severity score (1 to 3) from mild, moderate and severe – dermatologist diagnosed.

** Total IgE serum levels measured from blood samples by recruiting clinician.

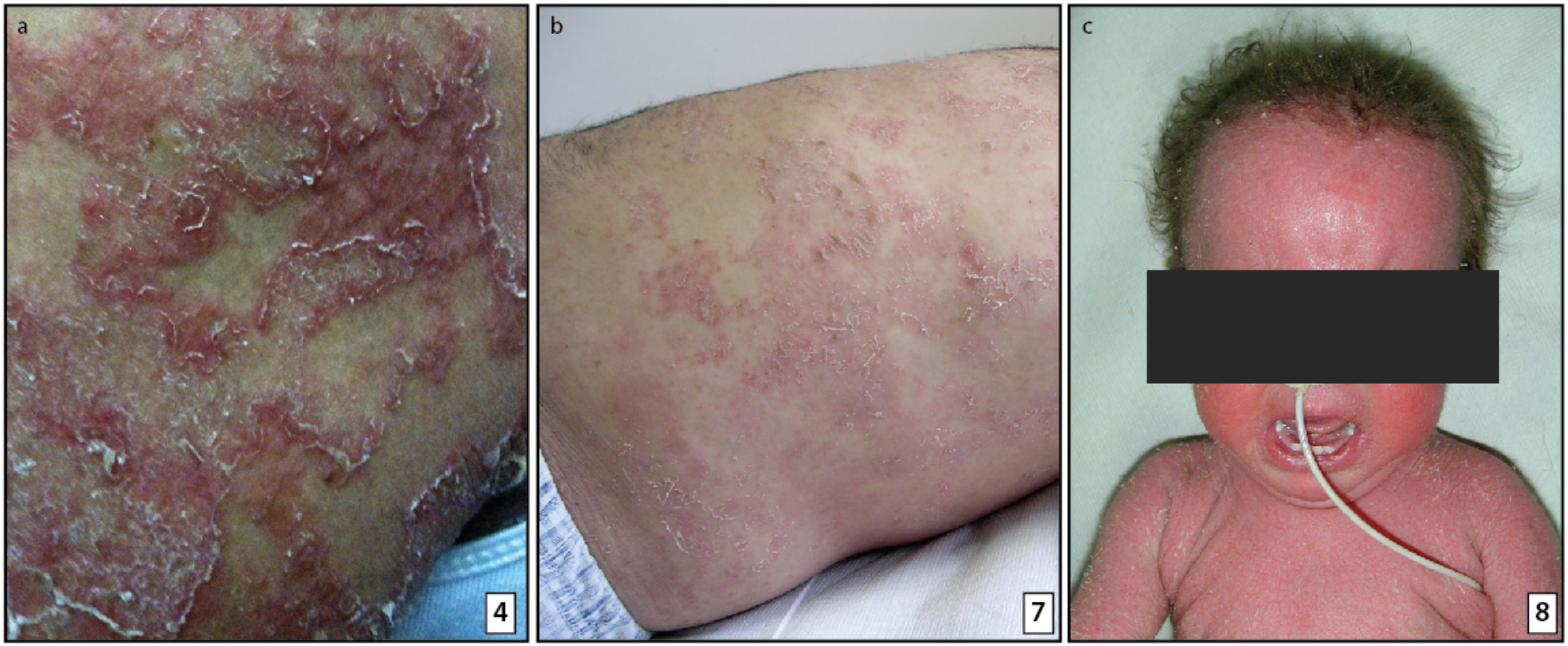



**Figure 4. 1 Clinical pictures of patients with Netherton Syndrome-like features.** Patient 4 (a; forearm), 7 (b; trunk) and 8 (c) showing extensive and severe ichthyosiform erythroderma, with widespread fine scaling as well as trichorrhexis invaginata (c).

| Demographic and clinical features | Paediatric Irish AD cases | Paediatric Irish AD cases with IgE > 10,000 (kU/L)* | South African AD cases | Northern Irish BKC cases |
|---|---|---|---|---|
| Total number (Females/Males) | 36 (25/11) | 7 (2/5) | 90 (50/40) | 55 (26/29) |
| Mean age (age range) (years) | 9 (1.6 to 15.8) | 6.8 (3.05 to 9.8) | 14 (5 to 49) | 53 (14 to 80) |
| Eczema severity score mean (range) | 13.48 (6 to 15)** | 13.86 (9 to 15)** | 10.5 (3 to 15)** | NA |

**Table 4. 2 Discovery cohorts of patients fully sequenced for *TMEM79*.** The patients were selected for their atopic phenotype combined with the absence of the 4 most common filaggrin mutations (R501X, 2282del4, R2447X and S3247X). AD, atopic dermatitis; BKC, blepharokeratoconjunctivitis.

* Total IgE concentrations measured using an ImmunoPhadia ELISA assay (Thermo Fisher Scientific, Uppsala, Sweden). Normal limits are between 120-150.

**Scored with Nottingham Eczema Severity Score (Emerson et al., 2000)

### 4.2.2. Sanger sequencing of *TMEM79*

*TMEM79* was screened using primers and conditions described in Chapter 2.2.2 (page 48) and Table 3.1 (page 76). All exons and intron/exon boundaries were sequenced. Amplification primers used cover all 3 variants that were discussed in Chapter 3.3.1 (page 88).

### 4.2.3. TaqMan® allelic discrimination assays

All cases and population controls were screened using TaqMan® allelic discrimination genotyping assays (Life Technologies Corporation, Applied Biosystems, California, USA) with a predesigned assay for SNV rs6684514 (catalog number C_25986870_10). For other SNVs that had no commercially available TaqMan® probes, custom design probes were made (Table 4. 3). 10 ng of template DNA was amplified in a 10 µL reaction with Perfecta® qPCR ToughMix ROX (Quanta, BioSciences Inc., Maryland, USA). The cycle conditions used were: 95°C for 10 mins; 40 cycles of 92°C for 15 seconds and 60°C for 1 min. PCR reactions were analyzed using a 7900HT Fast Real-Time PCR System and data analyzed with the 7900HT Sequence Detection System version 2.4 software (Life Technologies, Applied Biosystems, CA, USA).

| dbSNP ID | cDNA change | Protein change | F primer | R primer | Report 1 sequence (VIC) | Reporter 2 sequence (FAM) |
|---|---|---|---|---|---|---|
| rs112826258 | c.G92A | R31Q | GGATTCGGCCCCACCTT | AATGGCCGGCAGCC | AATGGCCGGCAGCC | AATGGCCAGCAGCC |
| rs75902624 | c.T310A | S104T | CTTCGGGCAGCTCCTCTAG | CCCCTGAGTCAGAACC | CCCCTGAGTCAGAACC | CCCTGAGACAGAACC |
| rs202080235 | c.A1232T | M354L | GGCTCCACCACGAACATGT | AGTAGAGGTTCCACTCAGC AT | AGTAGAGGTTCCACTCAGC AT | AGGTTCCACACAGCAT |

**Table 4. 3 Custom-design probes for allelic discrimination of *TMEM79* SNVs in the Xhosa population.** The bases highlighted in red show the base change within each probe for specific signal detection.

### 4.2.4. Statistical analysis

To test for the association of *TMEM79* SNVs with AD, logistics regression was used to calculate the odds ratios for each SNV, by comparing allelic frequencies between patient cohorts and population-matched controls. Power calculations were performed to calculate whether sample sizes were sufficient. All analyses were performed using the Sigmaplot™ v12.5 statistical package (Systat Software Inc., London, United Kingdom).

## 4.3. RESULTS

### 4.3.1. Netherton syndrome-like patients

The *TMEM79* gene was sequenced in 9 patients with phenotypes that were similar to Netherton Syndrome, but who were negative for any *SPINK5* mutations. These patients came from a variety of genetic backgrounds (Table 4. 1).

5 SNVs within *TMEM79* were identified (Table 4. 4), with only one coding for a protein change, rs6684514 (V147M). The odds ratio for this SNV could not be calculated as the patients were all from different ancestries, and population-matched controls could not be obtained.

| Chromosome 1 Position | 156,255,456 | 156255833 | 156,256,000 | 156,261,491 | 156,261,955 |
|---|---|---|---|---|---|
| mRNA position | 610 | - | - | 1458 | 1922 |
| dbSNP Reference No. | rs6684514 | rs2842883 | rs6679145 | rs3795728 | - |
| dbSNP MAF | 0.2115 | 0.0183 | 0.2143 | 0.2280 | - |
| Wildtype Allele | G | G | T | G | C |
| | V147M | intron | intron | 3' UTR | 3' UTR |
| 1 | WT | HOM C | WT | WT | WT |
| 2 | WT | HOM C | WT | HET C | WT |
| 3 | HET A | HOM C | HET C | HET C | WT |
| 4 | WT | HOM C | WT | WT | WT |
| 5 | HET A | HOM C | HET C | HET C | WT |
| 6 | WT | HOM C | WT | HOM C | HET T |
| 7 | WT | HOM C | WT | HET C | WT |
| 8 | WT | HOM C | WT | HET C | WT |
| 9 | HOM A | HOM C | HOM C | WT | WT |

**Table 4. 4 *TMEM79* SNVs found in patients displaying Netherton Syndrome-like phenotype.** Only 1 variant within the protein coding region, V147M (rs6684514).

### 4.3.2. Irish AD population

The *TMEM79* gene was fully sequenced in 43 Irish patients with AD including 7 with high levels of IgE (> 10,000 kU/L) where the 4 common filaggrin mutations had already been excluded. A total of 8 SNVs were identified (Table 4. 6), with only 1 SNV, rs6684514 (V147M) causing a missense mutation (Figure 4. 2a, page 150). In order to determine whether rs6684514 associated with AD risk, case-control analysis was carried out using 724 Irish AD patients and 1905 population-matched controls. The allelic frequency (Table 4. 5) showed that the minor allele (A) frequency (MAF) was the same in Irish AD cases and population-matched controls. Rs6684514 has no association with AD (P = 0.129; OR = 1.027, 95% CI = 0.900 – 1.172), indicating no significant association of rs6684514 with AD among the Irish population. The alpha value for the performed test was, 1.00, indicating that the population size is sufficient for this study.

| | rs6684514 (V147M) | | MAF | Odds ratio | P value | 95% CI |
|---|---|---|---|---|---|---|
| | Minor allele (A) | Major allele (G) | | | | |
| Irish AD | 437 | 1017 | 0.301 | 1.027 | 0.129 | 0.900 - 1.172 |
| Irish controls | 1127 | 2693 | 0.295 | | | |

**Table 4. 5 Logistics regression analysis of SNV rs6684514 in Irish AD.** No significant association was seen with V147M (rs6684514) within the Irish population.

| Chromosome 1 position | 156,253,031 | 156,255,456 | 156,255,748 | 156,255,833 | 156,256,000 | 156,261,491 | 156,261,754 | 156,261,955 |
|---|---|---|---|---|---|---|---|---|
| mRNA position | -- | 610 | -- | -- | -- | 1458 | -- | -- |
| dbSNP Reference No. | rs28372828 | rs6684514 | -- | rs2842883 | rs6679145 | rs3795728 | -- | -- |
| dbSNP MAF | 0.2704 | 0.2115 | -- | 0.0183 | 0.2143 | 0.228 | -- | -- |
| Base change | G > C | G > A | T >C | G > C | T >C | G >C | C > G | C >T |
| | 5' near gene | V147M | intron | intron | intron | 3' UTR | 3' UTR | 3' UTR |
| Irish AD patient genotypes: | | | | | | | | |
| Wildtype | 14 | 15 | 33 | 6 | 18 | 25 | 32 | 32 |
| Heterozygous | 20 | 18 | 0 | 0 | 16 | 8 | 1 | 1 |
| Homozygous | 2 | 3 | 3 | 27 | 1 | 0 | 0 | 0 |
| Irish AD patients (IgE > 10,000 kU/L) genotypes | | | | | | | | |
| Wildtype | 3 | 2 | 5 | 0 | 2 | 4 | 7 | 7 |
| Heterozygous | 2 | 3 | 0 | 0 | 3 | 2 | 0 | 0 |
| Homozygous | 2 | 2 | 2 | 7 | 2 | 1 | 0 | 0 |

**Table 4.6 *TMEM79* SNPs found in Irish AD population.** AD patients and AD patients with IgE > 10,000 (kU/L) were Sanger sequenced and only 1 variant was identified within the coding region, V147M (rs6684514).

### 4.3.3. Blepharokeratoconjunctivitis (BKC) patients

The *TMEM79* gene was sequenced in 30 Irish patients diagnosed with blepharokeratoconjunctivitis (BKC). The same spectrum of SNVs were found as in the Irish AD patients with rs66844514 being the only SNV found in the coding region. The allelic frequencies of rs6684514, showed no association with BKC (P = 0.0622; OR = 1.086, 95% CI = 0.702 – 1.682), indicating that the minor allele (A) does not affect BKC in the Irish population (Table 4. 7). The alpha value, 1.000, indicates power of the test is significant for this study.

| | rs6684514 (V147M) | | MAF | Odds ratio | P value | 95% CI |
|---|---|---|---|---|---|---|
| | Minor allele (A) | Major allele (G) | | | | |
| Irish BKC | 30 | 66 | 0.313 | 1.086 | 0.0622 | 0.702 – 1.682 |
| Irish controls | 1127 | 2693 | 0.295 | | | |

**Table 4. 7 Logistics regression analysis of Irish blepharokeratoconjunctivitis (BKC) patients.** The minor allele (A) rs6684514 showed no association with BKC within the Irish population.

### 4.3.4. Xhosa AD population

90 AD patients from the Xhosa population of South Africa had previously been screened as wildtype for all of the 4 common European filaggrin mutations. Included in this cohort were 31 patients in which the entire FLG gene had been sequenced but no loss-of-function mutations were identified. A pilot study of 31 Xhosa patients were fully screened for SNVs within the *TMEM79* gene, and a total of 11 SNVs (Table 4. 8) were identified with 4 occurring within the coding regions (Figure 4. 2). Case control analysis was carried out using 90 Xhosa AD patients and 82 Xhosa South African controls.

Using chi2 analysis, rs112826258 (R31Q) and rs75902624 (S104T) were not associated with AD as the P values were above 0.05 (Table 4. 9). Rs6684514 (V147M) is not associated with AD in the Xhosa population (P = 0.0152; OR = 0.921, 95% CI = 0.511 – 1.660). The raw data suggests that the minor allele (A) may play a protective role in the Xhosa AD patients (Table 4. 9) since it is more common in the controls than cases. The alpha power of the three performed tests above were between 1.000 showing that the sample population size is sufficient.

Novel SNV, M354L, was identified in two individuals from the Xhosa AD cohort and was not found in any of the South African controls. Unfortunately, chi2 analysis could not be performed since the expected frequency is less than 1, the 95% CI and the power of the performed test was 0.047 (less than 0.800), suggesting that the sample population should be increased in order to be able to detect a difference between cases and controls.

| Chromosome 1 Position | 156,252,734 | 156,255,109 | 156,255,327 | 156,255,456 | 156,255,635 | 156,255,833 | 156,256,000 | 156,261,153 | 156,261,265 | 156,261,399 | 156,261,491 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| mRNA position | -- | 263 | 481 | 610 | 789 | -- | -- | -- | 1232 | 1366 | 1458 |
| dbSNP Reference No. | rs116604461 | rs112826258 | rs75902624 | Rs6684514 | rs36114587 | rs2842883 | rs6679145 | rs7532219 | -- | rs113763043 | rs3795728 |
| dbSNP MAF | 0.126 | 0.005 | 0.0096 | 0.2115 | 0.0256 | 0.0183 | 0.2143 | 0.0041 | -- | 0.0046 | 0.228 |
| Base change | C > G | G > A | T > A | G > A | C > G | G > C | T > C | T > A | A > T | C > T | G > C |
| | 5' near gene | R31Q | S104T | V147M | A206 | intron | intron | intron | M354L | 3' UTR | 3' UTR |
| Xhosa AD patient genotypes: | | | | | | | | | | | |
| Wildtype | 25 | 21 | 25 | 20 | 24 | 0 | 20 | 20 | 25 | 20 | 16 |
| Heterozygous | 2 | 5 | 2 | 6 | 3 | 1 | 6 | 6 | 2 | 6 | 11 |
| Homozygous | 0 | 1 | 0 | 1 | 0 | 26 | 1 | 1 | 0 | 1 | 0 |

**Table 4. 8 *TMEM79* SNVs found in this Xhosa AD population using Sanger sequencing.** In a discovery cohort of 31 Xhosa patients, 4 SNVs were identified within the protein coding region: R31Q, S104T, V147M and M354L.

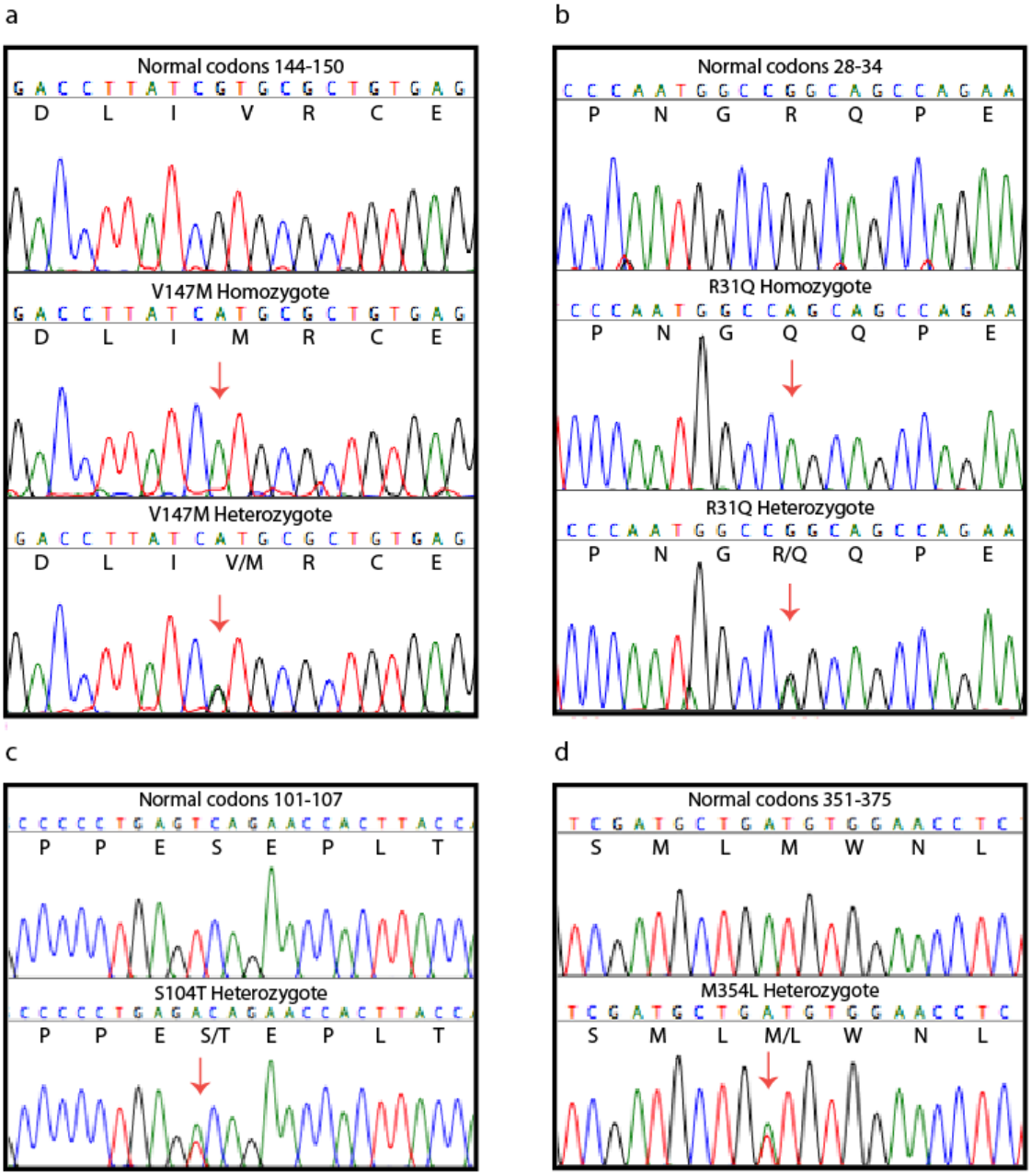



**Figure 4. 2 Sequencing chromatograms showing SNVs within the coding region of *TMEM79.*** V147M (a) in all populations, and R31Q, S104T and M354L in the Xhosa AD populations (b, c, d).

| | Minor allele | Major allele | MAF | Odds ratio | P value | 95% CI |
|---|---|---|---|---|---|---|
| rs112826258 (R31Q) | | | | | | |
| Xhosa AD | 24 | 156 | 0.133 | 0.817 | 0.61 | 0.448 - 1.488 |
| South african controls | 26 | 138 | 0.159 | | | |
| rs75902624 (S104T) | | | | | | |
| Xhosa AD | 6 | 174 | 0.033 | 1.851 | 0.59 | 0.455 – 7.522 |
| South african controls | 3 | 161 | 0.018 | | | |
| rs6684514 (V147M) | | | | | | |
| Xhosa AD | 28 | 152 | 0.156 | 0.658 | 0.18 | 0.379 – 1.141 |
| South african controls | 35 | 125 | 0.219 | | | |
| M354L | | | | | | |
| Xhosa AD | 2 | 176 | 0.011 | -- | -- | -- |
| South african controls | 0 | 154 | 0 | | | |

**Table 4. 9 Logistic regression analysis of SNVs identified within the Xhosa population.** Case-control analysis shows that rs112826258, rs75902624 and rs6684514 are not associated with AD within the Xhosa population. Chi2 analysis could not be performed for novel SNP M354L.

## 4.4. DISCUSSION

All exons of the *TMEM79* gene were sequenced in several discovery cohorts to identify any pathogenic variants. Besides having an overall non-specific skin inflammatory phenotype, the *Matt*$^{ma/ma}$ mouse also suffers from ocular and lung inflammation. Patients suffering from Netherton syndrome-like symptoms appear to most closely resemble the *Matt*$^{ma/ma}$ phenotype with eczema, high IgE levels, ocular inflammation, inflammatory lung condition and abnormal hair defects. Consequently, we sequenced these patients first, but the only coding variant found within these patients was the SNV, rs6684514. There was no association between the allelic frequencies of the minor allele with the severity of the patients' phenotype. The position of rs6684514 within the N-terminal region of the TMEM79 protein (Figure 4. 3) causes an amino acid substitution from valine to methione at position 147, but continues to retain the non-polar properties of the original amino acid. Rs6684514 was reported as having a MAF of 0.21 in a Genome Wide Association Study (GWAS) and had an association with erythrocyte-related traits in ~14,700 Japanese individuals ($P = 3.39 \times 10^{-9}$) (Kamatani et al., 2010). It is possible that TMEM79 may be affiliated with the lipid bilayer of the erythrocyte, which has been shown to be associated with multiple membrane proteins with a range of functions including, transport of molecules, adhesive properties and maintaining structural integrity. This SNV was also reported in the Exome Variant Server (EVS) (NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [July, 2013]) as having a MAF of 0.23 within a normal population made up of European American and African American samples, which was similar to the allele frequencies seen in our control cohorts. Rs6684514 was also present in the 1000 Genomes Project ((URL: www.1000genomes.org)

[July, 2013]), with a MAF of 0.211. With this SNV being present in 20% of the general population, it suggests that this SNV may not be pathogenic.

Have excluded *TMEM79* as a candidate for Netherton syndrome-like disease, we were naturally interested to find out if there was any association between variants in *TMEM79* with atopic dermatitis since the *matted* mouse is characterized by elevated levels of IgE and a spontaneous dermatitis phenotype.

The Xhosa AD patient cohort was screened for SNVs in *TMEM79* and a total of 4 missense variants were identified. This cohort had been found to have an absence of the common European *FLG* null mutations as well as any population-specific *FLG* mutations, and is a prime candidate for the involvement of other epidermal proteins in the pathogenesis of AD. The missense SNV rs6684514, which was also found in the Irish AD cohort, was identified as having a protective effect within the Xhosa AD cohort. This was contrary to what was seen in the Irish AD cohort, but may be attributed to population differences. Missense variant rs112826258 (R31Q) shows a similar odds ratio with rs6684514 (V147M) within the Xhosa population, suggesting that these two SNVs might be weakly associated in a protective manner (odds ratio less than 1). However, p values for rs112826258 (R31Q) and rs75902624 (S104T) are above 0.05, indicating that there is no association between these two SNVs and AD in the Xhosa population. The two SNVs rs112826258 and rs75902624 were present at very low frequencies within NCBI dbSNP, EVS and 1000 Genomes, with a MAF of less than 0.01 for both. The novel missense variant, M354L, identified within the Xhosa AD cohort was not present in the control population and interestingly, was not present within NCBI dbSNP Database, Exome Variant Server and 1000 Genomes Project, suggesting that this SNV is rare. At the same position however, a SNV rs202080325 (M354T) was identified in the CLINSEQ population (European descent) from NCBI dbSNP

Database and 1000 Genomes, but without any MAFs. As the number of affected individuals with this allele was small, case-control analysis with the South African control population could not be done. The two Xhosa AD patients with the SNV M354L were clinically diagnosed with moderate-severe AD severity, but records do not show any other noticeable clinical phenotype. Rs112826258 (R31Q), rs75902624 (S104T) and rs6684514 (V147M) are located within the N- terminal domain, while the novel variant M354L is within the fifth transmembrane region (Figure 4. 3). The change of a polar neutral methionine to leucine, which contains a hydrophobic side chain, may alter its' functional properties by causing structural changes within the fifth transmembrane domain of TMEM79. More investigations within various populations need to be done in order to clarify if this variant might have a functional impact on TMEM79 protein properties.

The Irish AD and Irish BKC patients that were sequenced in this chapter were selected as having a similar inflammatory phenotype (either epidermal or ocular) to the *Matt*$^{ma/ma}$ animal, and were all wildtype for the 4 most common *FLG* null mutations found in Europeans. Only one SNV, rs6684514, was identified as causing an amino acid change in TMEM79 but case-control analysis showed that it was not associated with either AD or BKC in the Irish population.

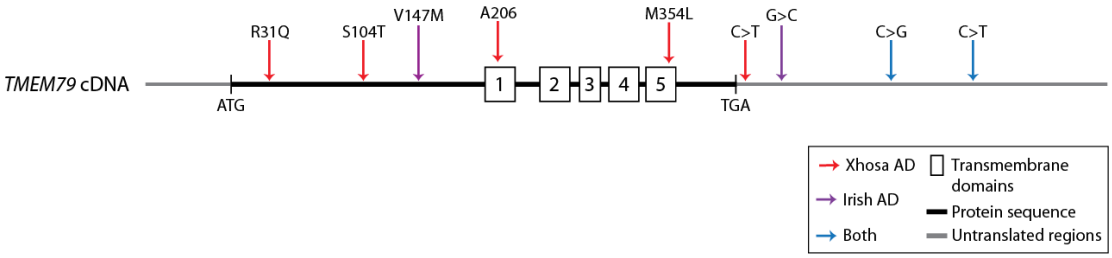



**Figure 4. 3 Location of SNVs within *TMEM79* cDNA**. Untranslated regions are seen in grey and protein coding sequences in black, with transmembrane domains shown as boxes.

GWAS are particularly useful in confirming associations with complex trait diseases, which do not obey standard Mendelian patterns of inheritance. Complex diseases are caused by a combination of genetic, environmental and lifestyle factors, and although a genetic predisposition means an increased susceptibility to developing disease, it does not necessarily follow that the person is destined to develop the disease. Ideally, a GWAS case-control study would require 2,000 or more cases and 3,000 ethinically-matched unaffected controls. This method of identifying susceptibility loci for complex skin diseases has so far been used for 10 skin diseases including systemic lupid erythematosus, vitiligo, leprosy, AD, male-pattern baldness, keloid, sarcoidosis, basal cell carcinoma, melanoma and cutaneous nevi (Zhang, 2012). In terms of AD, two GWAS in European and Chinese Han populations have been carried out which identified four susceptibility loci in AD cases: 1q21.3, 5q22.1, 11q13.5 and 20q13.33 (Esparza-Gordillo et al., 2009; Sun et al., 2011). Linkage disequilibrium was expected with *FLG* at 1q21, but within the European population, a risk allele (rs7927894) was identified for AD as well as for Crohn's disease (Barrett et al., 2008), and these two disorders share many inflammatory characteristics. New susceptibility loci identifying other candidate genes associated with the epidermal barrier and immune response suggests that there may be other mechanisms involved in the pathogenesis of AD. In case-control GWAS, it may sometimes be difficult to spot rare variants especially for complex diseases with locus/allelic heterogeneity, if they are not present in multiple samples or if there is low coverage even in large cohorts and may not be picked up during the variant calling step. This might occur if the program algorithms for variant-calling tools aren't programmed for sensitivity towards such variation (Goldstein et al., 2013).

In this chapter, a small group of 9 NS-like patients sharing a similar phenotype were sequenced for TMEM79. They are not good candidates for performing GWAS due to their small cohort size, heterogeneity of clinical phenotype and lack of ethnically matched controls due to mixed ancestries. It would be a better idea to perform whole exome sequencing on this group of patients, if ethical approval and appropriate normal controls could be obtained, as well as from non-affected family members if possible.

In the GWAS studies of AD described above, there was no significant association with the TMEM79 locus. This suggests that TMEM79 is not a genetic factor contributing towards AD and could explain why no association between variants in TMEM79 and AD was found. However, it is possible that TMEM79 may be indirectly associated with AD, perhaps via protein expression levels for example, akin to the down-regulation of filaggrin expression seen in AD patients without any FLG null mutations.

Whole transcriptome sequencing of patients with prostate cancer identified a unique RNA chimera of SMG5-TMEM79 in cancer tissues that was not seen in benign tissues, suggesting that this chimeric RNA could act as a potential biomarker. SMG5 is a transcript located immediately upstream of TMEM79. In cancerous tissues, the chimeric RNA found consisted of part of the third intron of TMEM79 spliced with two-thirds of SMG5 (Kannan et al., 2011). This was interesting to note, as I have showed that *TMEM79* transcript variant 1b is highly expressed in prostate tissues. In another study concerning asymptomatic left ventricular dysfunction (ALVD), which is the preclinical stage of heart failure, 7 genes were identified as being associated with ALVD using white blood cell gene expression profiling by microarrays. TMEM79 was one of these genes and showed a decreased expression in ALVD patients, and together with the other 6 genes,

have been shown to be an accurate diagnostic predictor for ALVD (Smih et al., 2011). Overall, these studies indicate as yet uncharacterised functions for TMEM79 outwith the context of the skin.

The screening of *TMEM79* within the relatively small Irish and Xhosa populations was performed to identify any potentially interesting SNVs that might help to better understand TMEM79 protein function within the skin. Since only one variant appeared to show significance within the discovery cohorts, further genotyping of rs6684514 was then carried out in other AD cohorts (England, UK, Ireland, Germany and Scotland) and the data is presented in Appendix VIII. I was involved in the screening of the English adult severe AD cohort, UK AD pediatric cases, Irish pediatric AD cases and their respective population controls. In a collaborative effort with various other institutes in an effort to screen different allergic patient populations, the missense mutation rs6684514, was further investigated for any association with AD (Appendix VIII). 4,245 AD cases from England, Scotland, Ireland, and Germany were combined and compared in a meta-analysis with 10,558 population-matched controls. The minor allele (A) appears to be protective for AD in all populations apart from the Irish cohort, which could be due to population or phenotypic differences. Figure 1 of Appendix VIII shows a meta-analysis of all the data which shows that rs6684514 has a small but significant effect on AD risk (P= 0.078; OR = 0.91, 95% CI = 0.86 – 0.96). Rs6684514, appears to be associated with AD risk in a small but significant manner across different AD populations.

In conclusion, if cost and time were not an issue, it would be more productive to sequence the Irish and Xhosa patient cohorts by whole exome analysis instead of performing traditional Sanger sequencing, as this might identify other potential

gene associations. The work carried out in this chapter was performed before the evidence of TMEM79 associating with LG was uncovered, and at that point, the most likely discovery cohorts based on phenotypic similarity were shortlisted for screening. Based on other recent publications, it is interesting to note that TMEM79 appears to be involved with conditions not relating to epidermal disease, and this could involving the other variants of TMEM79 which I have been shown to be expressed in other tissues.

# CHAPTER 5

# OVERALL CONCLUSION AND FUTURE WORK

**OVERALL CONCLUSION**

The *Matt*$^{ma/ma}$*Flg*$^{ft/ft}$ double mutant (DM) mouse first arose spontaneously on the background of the *matted* mouse, and the DM animal has been used as a model in skin barrier studies for human atopic dermatitis (AD) because it is filaggrin-deficient. The mutation responsible for the flaky-tail mouse phenotype had been identified earlier as a frameshift mutation in the *Flg* gene (Fallon et al., 2009) In this thesis, I set out to identify the gene responsible for the atopic appearance and hair shaft defect of the *Matt*$^{ma/ma}$ mouse. Using data obtained from whole transcriptome sequencing and homozygozity mapping strategies, a novel gene was identified, *Tmem79*. This gene is encoded by 4 exons, and translates into a 5 transmembrane domain protein. *Matt*$^{ma/ma}$ harbors a spontaneous mutation, c.840C>G or p.Y280*, within the third transmembrane region of *Tmem79*. Multiple sequence alignment shows that TMEM79 is highly conserved across species, and protein prediction software analysis suggests distant homology to members of the MAPEG family. The results shown in Chapter 2 have been submitted to the Journal of Allergy and Clinical Immunology in July 2013 and have been accepted for publication in November 2013.

During the progression of this project, a second mutant mouse was identified as having a missense mutation in *Tmem79* (MGI Ref ID J:178673). The Jackson Laboratory reported this mutant in 2011, known as 129;FVB-Tmem79$^{m1J}$/GrsrJ, and was described as having an abnormal hair morphology with kinks, a sparse coat and some adults display mild irritation around the eyes (Figure 5. 1). Unlike *Matt*$^{ma/ma}$ this mouse harbors a missense mutation, p.G214A which occurs in the first transmembrane domain (Figure 5. 2). In comparison, the *Matt*$^{ma/ma}$ mouse which has a truncated TMEM79 due to the nonsense mutation, p.Y280*, has a more severe inflammatory phenotype.

**Tmem79$^{m1J}$/Tmem79$^{m1J}$**
129;FVB-Tmem79$^{m1J}$/GrsrJ

**integument phenotype**
abnormal coat/hair pigmentation
- homozygotes on an albino background appear off-white (MGI Ref ID J:178673)

abnormal skin condition
- some adults display mild irritation around the eyes (MGI Ref ID J:178673)

abnormal zigzag hair morphology (MGI Ref ID J:178673)

increased curvature of zigzag hairs
- one to two additional kinks are found in the zigzag hairs and the ends are not straight (MGI Ref ID J:178673)

shiny fur
- in young homozygotes the coat can appear slightly shiny (MGI Ref ID J:178673)

sparse hair
- the coat appears more sparse than normal but not severely sparse (MGI Ref ID J:178673)

**pigmentation phenotype**
abnormal coat/hair pigmentation
- homozygotes on an albino background appear off-white (MGI Ref ID J:178673)

**Figure 5. 1 Phenotype of the 129;FVB-Tmem79$^{m1j}$/GrsrJ mouse mutant.** A mouse mutant with a reported missense mutation in *Tmem79*, p.G214A. Image screen captured from JAX® Mice Database.

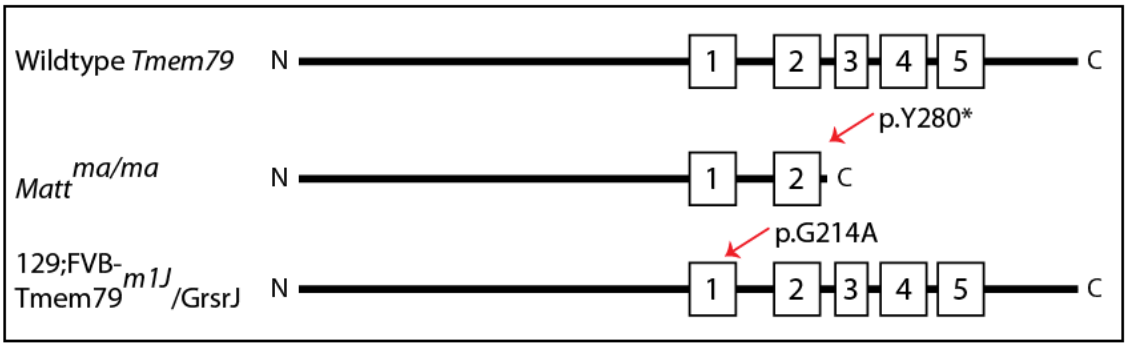



**Figure 5. 2 Location of murine mutations within TMEM79.** *Matt$^{ma/ma}$* would result in a prematurely truncated protein and 129;FVB-Tmem79$^{m1J}$/GrsrJ would still express the full-length protein.

In humans, this novel gene is located within the EDC and is expressed only in the upper granular layers of the epidermis where it is localized to the membranes of lamellar granules (LG). These are branched, tubular structures originating from the trans-Golgi network (TGN) and contain various cargoes meant for secretion at the apical surface of the topmost granular cells (Ishida-Yamamoto et al., 2004). These bulbous structures have been shown to contain proteins such as cornedesmosin (CDSN) and various hydrolytic enzymes such as kallikrein 5,7 and 8 (KLK5, KLK7 and KLK8), cathepsin D (CatD) and glucosylceramides (GlcCer)

(Freinkel and Traczyk, 1983; Freinkel and Traczyk, 1985; Grayson et al., 1985; Ishida-Yamamoto et al., 2004; Serre et al., 1991). Immunofluorescent imaging of TMEM79, CDSN and KLK8 show granular aggregates, whilst transmission electron microscopy shows that these molecules are housed within distinct tubular reticulate structures. The contents of LG bodies must aggregate at specific storage locations before they are required, but the mechanisms and timing for the release of LG cargoes are still unknown. There is evidence to suggest that CDSN is expressed the earliest, followed by GlcCer and finally by KLK7 and KLK8 (Ishida-Yamamoto et al., 2004).

Through the course of this work, a collaborative partner laboratory independently identified and confirmed that the *Tmem79* gene is indeed responsible for the matted phenotype of the *Matt*$^{ma/ma}$ animal (Sasaki et al., 2013).They identified the *Tmem79* gene and mutation by performing backcrosses and applying target enrichment, followed by next generation sequencing. They then rescued the phenotype of the *Matt*$^{ma/ma}$ mouse by introducing a wildtype *Tmem79* transgene under the control of *Tmem79* promoter, and showed that in these mutant mice, TMEM79 was localized to the LG secretory system. Furthermore, they showed that *Matt*$^{ma/ma}$ mice have a decreased expression of LG-secreted proteins such as KLK5, KLK7 and LEKTI in comparison to wildtype mice and that cornified layers could be removed *en bloc* through epidermal tape-stripping experiments. An antibody was made against the N-terminus of TMEM79, and showed that the truncated form of TMEM79 was still detectable in the upper granular layers, suggesting that the C-terminus of the protein is functionally important. Their findings suggest that in *Matt*$^{ma/ma}$ mice, the impairment of TMEM79 leads to SC barrier dysfunction. The characterization of the *Matt*$^{ma/ma}$ mouse by our Japanese collaborators have produced results that correspond well with the work carried out in this thesis.

The role of TMEM79 could be as a labeling molecule for identification of a particular cargo within a LG, or an anchoring protein that aids in the transport process of LG or during fusion with the cells' plasma membrane (Figure 5. 3).

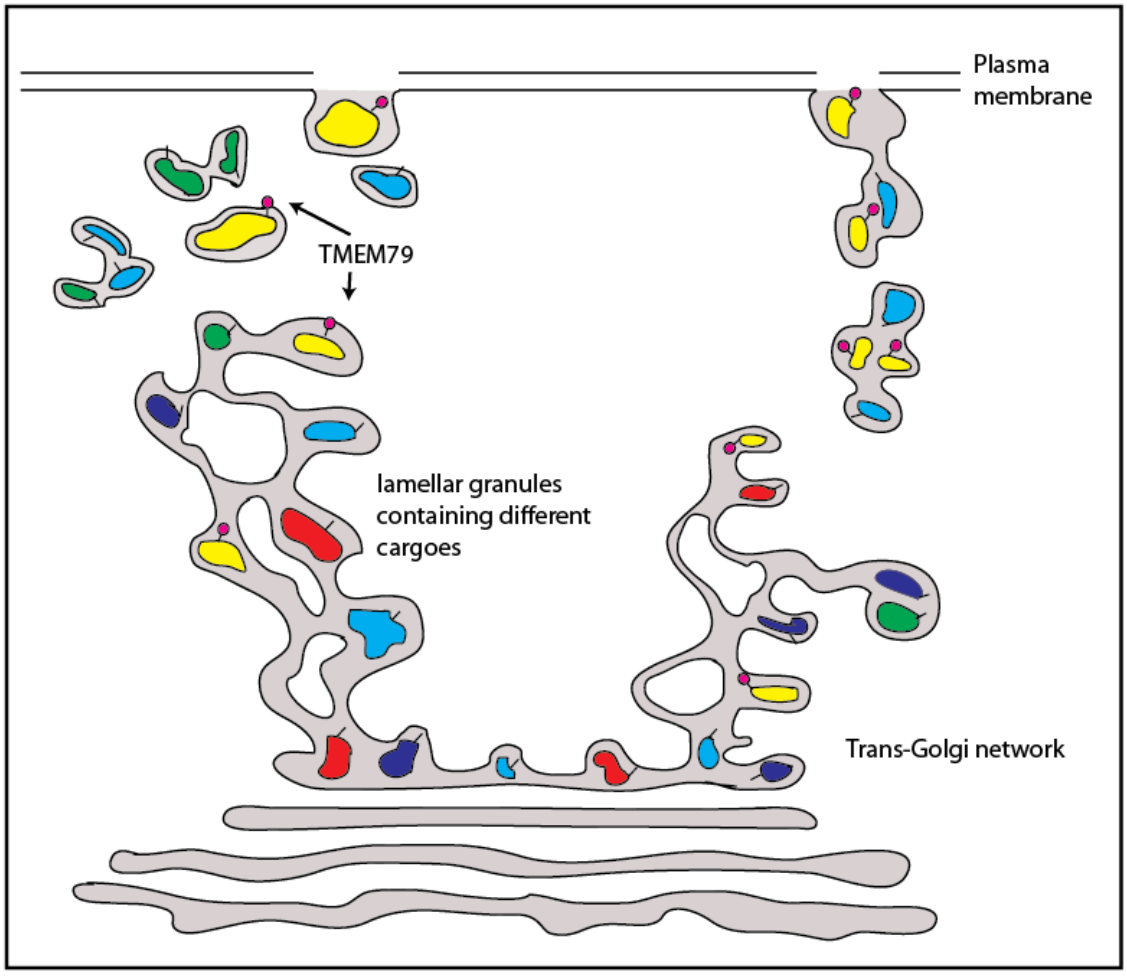



**Figure 5. 3 Cartoon showing the hypothetical role of TMEM79.** Lamellar granules (LG) are connected to the TGN and transport the contents of LGs individually for secretion at the apical face of the most superficial granulocytes.

The release of lipids is vital for the lipid envelope of the stratum corneum. Disruption of this envelope allows penetration of allergens causing an inflammatory response. Some diseases affected by an impaired epidermal lipid barrier have been detailed earlier in Chapter 1.4.2, page 33. To date, no human equivalent of the matted mutation has been identified in the populations that have been screened. The common SNV, rs6684514, appears to have a protective effect with AD. It was initially thought that blepharitis was a good candidate disease, as

both the *Matt*$^{ma/ma}$ and 129;FVB-Tmem79$^{m1J}$/GrsrJ mouse mutants display ocular inflammation, but no pathogenic mutations were identified in BKC patients. At the stage when screening for human disease association was carried out, the information regarding the involvement of TMEM79 with lamellar granules was still not discovered, hence, human disorders resembling the eczema-like inflammatory phenotype of the *Matted*$^{ma/ma}$ mouse were chosen as candidate groups. It is possible that the human diseases screened in this study were too mild, and that loss of TMEM79 protein may result in a phenotype more severe than AD in humans. According to the subcellular location of the TMEM79 protein within the membranes of LGs, a more accurate human clinical phenotype resulting from aberrant TMEM79 expression may resemble a form of lamellar ichthyosis disease.

**FUTURE WORK**

Double immuno-gold labeling experiments performed in this thesis have ruled out any association between CDSN and KLK8 with TMEM79-labelled LGs. More labeling experiments need to be performed in order to identify the contents of TMEM79-labelled secretory granules. Along with the known LG-contents (listed above), an interesting candidate to look at would be caveolin-1, which is a cholesterol-binding protein that is thought to be involved in LG assembly or trafficking in keratinocytes of the epidermis (Sando et al., 2003).

The commercial polyclonal antibody against TMEM79 that was used in the immunofluorescent and immunoelectron microscopy imaging experiments is directed against the C-terminus. It would be useful to develop an antibody against the N-terminus of the TMEM79 protein to allow the visualization of any truncated TMEM79 protein in the epidermis of the *Matt*$^{ma/ma}$ animal, and would also

overcome the difficulties of TMEM79 fusion constructs, where the epitope could be blocked by different tags. A traditional approach to identify potential TMEM79 partners would be to perform immunoprecipitation assays, but could be problematic due its transmembrane characteristics. Alternatively, a proximity ligation immunoassay or an affinity column of purified TMEM79 could be used to detect protein-protein interactions. These might also confirm if TMEM79 has any glutathione binding/activity as bioinformatics analysis has predicted.

Results obtained from overexpression assays performed with HaCaT cells and HPKs did not yield any conclusive results about the cellular location of TMEM79 since overexpressed TMEM79 aggregated and was disposed by autophagocytosis. Taking into account that TMEM79 is only expressed in keratinocytes of the upper granular layer, cell lines could perhaps have been cultured in low calcium conditions (0.02 – 0.1 mM) before stratifying in high calcium conditions (1.8 – 2.0 mM). Preliminary experiments could be carried out to observe for any endogenous TMEM79 expression under these conditions.

Another alternative would be to establish a stable keratinocyte cell line with a tagged TMEM79 construct, and this could be used in a 3D-culture to ensure keratinocyte stratification and TMEM79 expression. Preliminary experiments were carried out trying to detect endogenous TMEM79 in a 3D-culture, but none could be detected. This was probably due to having only very small amounts of sample, as well as an antibody that is not efficient at detecting proteins on an immunoblot. A way to accomplish this could be the use of a retroviral expression system with human primary keratinocytes, with the eventual aim of stratifying these stably transfected human primary keratinocytes in a 3D skin model.

The same could also be performed using mouse keratinocytes, and *Tmem79* could conceivably be knocked out using siRNA targeted against *Tmem79*, and observing any abnormalities in stratification afterwards. This might allow the focus to be on the fundamental mechanism of TMEM79 *in vitro*, without the secondary inflammatory effects that are so prevalent in the mouse model.

Next Generation Sequencing could be carried out for the Xhosa and Irish AD patients who did not have any identified *FLG* mutations. In theory, it would be easier to flag up any susceptibility loci and candidate genes using this NGS approach. As ethnically-matched controls for these populations already exist, the only requirement left would be to obtain informed consent from these patients to proceed.

Finally, it should be noted that *TMEM79* might be involved with other pathophysiological mechanisms as discussed previously (e.g. erythrocyte-related traits, ALVD and prostate cancer) and these could involve other *TMEM79* transcript variants. Disease pathology may appear different in humans with a *TMEM79* mutation than in the $Matt^{ma/ma}$ mouse. Similarly, the phenotype of a TMEM79 knockout mouse could also differ considerably from $Matt^{ma/ma}$ mice. Overall, there are many unanswered questions remaining, and much more work needs to be carried out to understand the functions of TMEM79.

# REFERENCES

Aho, S., K. Li, Y. Ryoo, C. McGee, A. Ishida-Yamamoto, J. Uitto, and J.F. Klement. 2004. Periplakin gene targeting reveals a constituent of the cornified cell envelope dispensable for normal mouse development. *Molecular and cellular biology*. 24:6410-6418.

Akiyama, M. 2010. ABCA12 mutations and autosomal recessive congenital ichthyosis: a review of genotype/phenotype correlations and of pathogenetic concepts. *Human mutation*. 31:1090-1096.

Akiyama, M., I. Matsuo, and H. Shimizu. 2002. Formation of cornified cell envelope in human hair follicle development. *The British journal of Dermatology*. 146:968-976.

Akiyama, M., Y. Sugiyama-Nakagiri, K. Sakai, J.R. McMillan, M. Goto, K. Arita, Y. Tsuji-Abe, N. Tabata, K. Matsuoka, R. Sasaki, D. Sawamura, and H. Shimizu. 2005. Mutations in lipid transporter ABCA12 in harlequin ichthyosis and functional recovery by corrective gene transfer. *The Journal of Clinical Investigation*. 115:1777-1784.

Al-Jassar, C., H. Bikker, M. Overduin, and M. Chidgey. 2013. Mechanistic Basis of Desmosome-Targeted Diseases. *Journal of molecular biology*.

Alibardi, L. 2004. Comparative aspects of the inner root sheath in adult and developing hairs of mammals in relation to the evolution of hairs. *Journal of Anatomy*. 205:179-200.

Alibardi, L., and N. Bernd. 2013. Immunolocalization of junctional proteins in human hairs indicates that the membrane complex stabilizes the inner root sheath while desmosomes contact the companion layer through specific keratins. *Acta Histochemica*.

Alonso, L., H. Okada, H.A. Pasolli, A. Wakeham, A.I. You-Ten, T.W. Mak, and E. Fuchs. 2005. Sgk3 links growth factor signaling to maintenance of progenitor cells in the hair follicle. *The Journal of Cell Biology*. 170:559-570.

Ammar, M., F. Bouazizi, R. Bouhaha, I. Zaraa, S. Kouidhi, S. Ourheni, C. Helms, N. Doss, R. Dhaoui, A. Ben Osman, A. Ben Ammar-El Gaaied, R. Marrakchi, M. Mokni, and C. Bouchlaka-Souissi. 2012. Association analysis of LCE3C-LCE3B deletion in Tunisian psoriatic population. *Archives of dermatological research*. 304:733-738.

Anders, S., and W. Huber. 2010. Differential expression analysis for sequence count data. *Genome Biology*. 11:R106.

Andl, T., K. Ahn, A. Kairo, E.Y. Chu, L. Wine-Lee, S.T. Reddy, N.J. Croft, J.A. Cebra-Thomas, D. Metzger, P. Chambon, K.M. Lyons, Y. Mishina, J.T. Seykora, E.B. Crenshaw, 3rd, and S.E. Millar. 2004. Epithelial Bmpr1a regulates differentiation and proliferation in postnatal hair follicles and is essential for tooth development. *Development*. 131:2257-2268.

Armstrong, D.K., K.E. McKenna, P.E. Purkis, K.J. Green, R.A. Eady, I.M. Leigh, and A.E. Hughes. 1999. Haploinsufficiency of desmoplakin causes a striate subtype of palmoplantar keratoderma. *Human molecular genetics*. 8:143-148.

Baker, R.E., and P.J. Murray. 2012. Understanding hair follicle cycling: a systems approach. *Current opinion in genetics & development*. 22:607-612.

Bao, L., F. Sachs, and G. Dahl. 2004. Connexins are mechanosensitive. *American journal of physiology. Cell physiology*. 287:C1389-1395.

Barrett, J.C., S. Hansoul, D.L. Nicolae, J.H. Cho, R.H. Duerr, J.D. Rioux, S.R. Brant, M.S. Silverberg, K.D. Taylor, M.M. Barmada, A. Bitton, T. Dassopoulos, L.W. Datta, T. Green, A.M. Griffiths, E.O. Kistner, M.T. Murtha, M.D. Regueiro, J.I. Rotter, L.P. Schumm, A.H. Steinhart, S.R. Targan, R.J. Xavier, C. Libioulle, C. Sandor, M. Lathrop, J. Belaiche, O. Dewit, I. Gut, S. Heath, D. Laukens, M. Mni, P. Rutgeerts,

A. Van Gossum, D. Zelenika, D. Franchimont, J.P. Hugot, M. de Vos, S. Vermeire, E. Louis, L.R. Cardon, C.A. Anderson, H. Drummond, E. Nimmo, T. Ahmad, N.J. Prescott, C.M. Onnie, S.A. Fisher, J. Marchini, J. Ghori, S. Bumpstead, R. Gwilliam, M. Tremelling, P. Deloukas, J. Mansfield, D. Jewell, J. Satsangi, C.G. Mathew, M. Parkes, M. Georges, and M.J. Daly. 2008. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature genetics*. 40:955-962.

Basler, E., M. Grompe, G. Parenti, J. Yates, and A. Ballabio. 1992. Identification of point mutations in the steroid sulfatase gene of three patients with X-linked ichthyosis. *American Journal of Human Genetics*. 50:483-491.

Behne, M.J., and J.M. Jensen. 2012. Calcium in epidermis. *Advances in experimental medicine and biology*. 740:945-953.

Behne, M.J., J.W. Meyer, K.M. Hanson, N.P. Barry, S. Murata, D. Crumrine, R.W. Clegg, E. Gratton, W.M. Holleran, P.M. Elias, and T.M. Mauro. 2002. NHE1 regulates the stratum corneum permeability barrier homeostasis. Microenvironment acidification assessed with fluorescence lifetime imaging. *The Journal of Biological Chemistry*. 277:47399-47406.

Bergboer, J.G., M. Umicevic-Mirkov, J. Fransen, M. den Heijer, B. Franke, P.L. van Riel, J. Schalkwijk, and M.J. Coenen. 2012. A replication study of the association between rheumatoid arthritis and deletion of the late cornified envelope genes LCE3B and LCE3C. *PloS one*. 7:e32045.

Bernard, B.A. 2012. The human hair follicle, a bistable organ? *Experimental Dermatology*. 21:401-403.

Bernard, D., B. Mehul, A. Thomas-Collignon, L. Simonetti, V. Remy, M.A. Bernard, and R. Schmidt. 2003. Analysis of proteins with caseinolytic activity in a human stratum corneum extract revealed a yet unidentified cysteine protease and identified the so-called "stratum corneum thiol protease" as cathepsin l2. *The Journal of Investigative Dermatology*. 120:592-600.

Bikle, D.D., and S. Pillai. 1993. Vitamin D, calcium, and epidermal differentiation. *Endocrine reviews*. 14:3-19.

Blanpain, C., W.E. Lowry, A. Geoghegan, L. Polak, and E. Fuchs. 2004. Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. *Cell*. 118:635-648.

Blaydon, D.C., D. Nitoiu, K.M. Eckl, R.M. Cabral, P. Bland, I. Hausser, D.A. van Heel, S. Rajpopat, J. Fischer, V. Oji, A. Zvulunov, H. Traupe, H.C. Hennies, and D.P. Kelsell. 2011. Mutations in CSTA, encoding Cystatin A, underlie exfoliative ichthyosis and reveal a role for this protease inhibitor in cell-cell adhesion. *American journal of human genetics*. 89:564-571.

Bonifas, J.M., J.W. Bare, M.A. Chen, M.K. Lee, C.A. Slater, L.A. Goldsmith, and E.H. Epstein, Jr. 1992. Linkage of the epidermolytic hyperkeratosis phenotype and the region of the type II keratin gene cluster on chromosome 12. *The Journal of Investigative Dermatology*. 99:524-527.

Boukamp, P., R.T. Petrussevska, D. Breitkreutz, J. Hornung, A. Markham, and N.E. Fusenig. 1988. Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. *The Journal of Cell Biology*. 106:761-771.

Bowden, P.E., J.L. Haley, A. Kansky, J.A. Rothnagel, D.O. Jones, and R.J. Turner. 1995. Mutation of a type II keratin gene (K6a) in pachyonychia congenita. *Nature Genetics*. 10:363-365.

Burkitt, H.G., B. Young, and J.W. Heath. 1993. Wheater's Functional Histology, A Text and Colour Atlas. Churchill Livingstone, Medical Division of Longman Group UK Ltd, Edinburgh.

Cabral, A., A. Sayin, S. de Winter, D.F. Fischer, S. Pavel, and C. Backendorf. 2001. SPRR4, a novel cornified envelope precursor: UV-dependent epidermal expression and selective incorporation into fragile envelopes. *Journal of cell science*. 114:3837-3843.

Candi, E., G. Melino, A. Lahm, R. Ceci, A. Rossi, I.G. Kim, B. Ciani, and P.M. Steinert. 1998a. Transglutaminase 1 mutations in lamellar ichthyosis. Loss of activity due to failure of activation by proteolytic processing. *The Journal of Biological Chemistry*. 273:13693-13702.

Candi, E., G. Melino, G. Mei, E. Tarcsa, S.I. Chung, L.N. Marekov, and P.M. Steinert. 1995. Biochemical, structural, and transglutaminase substrate properties of human loricrin, the major epidermal cornified cell envelope protein. *The Journal of Biological Chemistry*. 270:26382-26390.

Candi, E., S. Oddi, A. Terrinoni, A. Paradisi, M. Ranalli, A. Finazzi-Agro, and G. Melino. 2001. Transglutaminase 5 cross-links loricrin, involucrin, and small proline-rich proteins in vitro. *The Journal of Biological Chemistry*. 276:35014-35023.

Candi, E., R. Schmidt, and G. Melino. 2005. The cornified envelope: a model of cell death in the skin. *Nature Reviews: Molecular cell biology*. 6:328-340.

Candi, E., E. Tarcsa, J.J. Digiovanna, J.G. Compton, P.M. Elias, L.N. Marekov, and P.M. Steinert. 1998b. A highly conserved lysine residue on the head domain of type II keratins is essential for the attachment of keratin intermediate filaments to the cornified cell envelope through isopeptide crosslinking by transglutaminases. *Proceedings of the National Academy of Sciences of the United States of America*. 95:2067-2072.

Cassidy, A.J., M.A. van Steensel, P.M. Steijlen, M. van Geel, J. van der Velden, S.M. Morley, A. Terrinoni, G. Melino, E. Candi, and W.H. McLean. 2005. A homozygous missense mutation in TGM5 abolishes epidermal transglutaminase 5 activity and causes acral peeling skin syndrome. *American Journal of Human Genetics*. 77:909-917.

Chavanas, S., C. Bodemer, A. Rochat, D. Hamel-Teillac, M. Ali, A.D. Irvine, J.L. Bonafe, J. Wilkinson, A. Taieb, Y. Barrandon, J.I. Harper, Y. de Prost, and A. Hovnanian. 2000. Mutations in SPINK5, encoding a serine protease inhibitor, cause Netherton syndrome. *Nature Genetics*. 25:141-142.

Connolly, S.A., D.J. Landsburg, A. Carfi, D.C. Wiley, G.H. Cohen, and R.J. Eisenberg. 2003. Structure-based mutagenesis of herpes simplex virus glycoprotein D defines three critical regions at the gD-HveA/HVEM binding interface. *Journal of Virology*. 77:8127-8140.

Coulombe, P.A., M.L. Kerns, and E. Fuchs. 2009. Epidermolysis bullosa simplex: a paradigm for disorders of tissue fragility. *The Journal of clinical investigation*. 119:1784-1793.

Covello, S.P., A.D. Irvine, K.E. McKenna, C.S. Munro, N.C. Nevin, F.J. Smith, J. Uitto, and W.H. McLean. 1998. Mutations in keratin K9 in kindreds with epidermolytic palmoplantar keratoderma and epidemiology in Northern Ireland. *The Journal of Investigative Dermatology*. 111:1207-1209.

Dale, B.A., and E. Kam. 1993. Harlequin ichthyosis. Variability in expression and hypothesis for disease mechanism. *Archives of Dermatology*. 129:1471-1477.

Dale, B.A., R.B. Presland, S.P. Lewis, R.A. Underwood, and P. Fleckman. 1997. Transient expression of epidermal filaggrin in cultured cells causes collapse of

intermediate filament networks with alteration of cell shape and nuclear integrity. *The Journal of Investigative Dermatology*. 108:179-187.

Dannies, P.S. 2001. Concentrating hormones into secretory granules: layers of control. *Molecular and Cellular Endocrinology*. 177:87-93.

De Laurenzi, V., G.R. Rogers, D.J. Hamrock, L.N. Marekov, P.M. Steinert, J.G. Compton, N. Markova, and W.B. Rizzo. 1996. Sjogren-Larsson syndrome is caused by mutations in the fatty aldehyde dehydrogenase gene. *Nature Genetics*. 12:52-57.

Descargues, P., C. Deraison, C. Bonnart, M. Kreft, M. Kishibe, A. Ishida-Yamamoto, P. Elias, Y. Barrandon, G. Zambruno, A. Sonnenberg, and A. Hovnanian. 2005. Spink5-deficient mice mimic Netherton syndrome through degradation of desmoglein 1 by epidermal protease hyperactivity. *Nature genetics*. 37:56-65.

DiGiovanna, J.J., and S.J. Bale. 1994. Clinical heterogeneity in epidermolytic hyperkeratosis. *Archives of Dermatology*. 130:1026-1035.

Djian, P., K. Easley, and H. Green. 2000. Targeted ablation of the murine involucrin gene. *The Journal of Cell Biology*. 151:381-388.

Dong, W., M. Ryynanen, and J. Uitto. 1993. Identification of a leucine-to-proline mutation in the keratin 5 gene in a family with the generalized Kobner type of epidermolysis bullosa simplex. *Human Mutation*. 2:94-102.

Downing, D.T., R.W. Dose, and W. Abraham. 1993. Interaction between sphingosine and cholesteryl sulfate in epidermal lipids. *Journal of Lipid Research*. 34:563-569.

Downing, D.T., M.E. Stewart, P.W. Wertz, S.W. Colton, W. Abraham, and J.S. Strauss. 1987. Skin lipids: an update. *The Journal of Investigative Dermatology*. 88:2s-6s.

Drees, F., S. Pokutta, S. Yamada, W.J. Nelson, and W.I. Weis. 2005. Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly. *Cell*. 123:903-915.

Ebihara, L. 1996. Xenopus connexin38 forms hemi-gap-junctional channels in the nonjunctional plasma membrane of Xenopus oocytes. *Biophysical Journal*. 71:742-748.

Eckert, R.L., J.F. Crish, T. Efimova, S.R. Dashti, A. Deucher, F. Bone, G. Adhikary, G. Huang, R. Gopalakrishnan, and S. Balasubramanian. 2004. Regulation of involucrin gene expression. *The Journal of investigative dermatology*. 123:13-22.

Eckert, R.L., J.F. Crish, and N.A. Robinson. 1997. The epidermal keratinocyte as a model for the study of gene regulation and cell differentiation. *Physiological Reviews*. 77:397-424.

Eckert, R.L., and H. Green. 1986. Structure and evolution of the human involucrin gene. *Cell*. 46:583-589.

Eckhart, L., S. Lippens, E. Tschachler, and W. Declercq. 2013. Cell death by cornification. *Biochimica et biophysica acta*.

Eicher, E.M., and P.W. Lane. 1980. Assignment of LG XVI to chromosome 3 in the mouse. *Journal of Heredity*. 71:315-318.

Elias, P., S. Ahn, B. Brown, D. Crumrine, and K.R. Feingold. 2002. Origin of the epidermal calcium gradient: regulation by barrier status and role of active vs passive mechanisms. *The Journal of investigative dermatology*. 119:1269-1274.

Elias, P.M., P. Fritsch, and E.H. Epstein. 1977. Staphylococcal scalded skin syndrome. Clinical features, pathogenesis, and recent microbiological and biochemical developments. *Archives of Dermatology*. 113:207-219.

Emerson, R.M., C.R. Charman, and H.C. Williams. 2000. The Nottingham Eczema Severity Score: preliminary refinement of the Rajka and Langeland grading. *The British Journal of Dermatology*. 142:288-297.

Esparza-Gordillo, J., S. Weidinger, R. Folster-Holst, A. Bauerfeind, F. Ruschendorf, G. Patone, K. Rohde, I. Marenholz, F. Schulz, T. Kerscher, N. Hubner, U. Wahn, S. Schreiber, A. Franke, R. Vogler, S. Heath, H. Baurecht, N. Novak, E. Rodriguez, T. Illig, M.A. Lee-Kirsch, A. Ciechanowicz, M. Kurek, T. Piskackova, M. Macek, Y.A. Lee, and A. Ruether. 2009. A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nature genetics*. 41:596-601.

Fallon, P.G., T. Sasaki, A. Sandilands, L.E. Campbell, S.P. Saunders, N.E. Mangan, J.J. Callanan, H. Kawasaki, A. Shiohama, A. Kubo, J.P. Sundberg, R.B. Presland, P. Fleckman, N. Shimizu, J. Kudoh, A.D. Irvine, M. Amagai, and W.H. McLean. 2009. A homozygous frameshift mutation in the mouse Flg gene facilitates enhanced percutaneous allergen priming. *Nature Genetics*. 41:602-608.

Fischer, H., S. Szabo, J. Scherz, K. Jaeger, H. Rossiter, M. Buchberger, M. Ghannadan, M. Hermann, H.C. Theussl, D.J. Tobin, E.F. Wagner, E. Tschachler, and L. Eckhart. 2011. Essential role of the keratinocyte-specific endonuclease DNase1L2 in the removal of nuclear DNA from hair and nails. *The Journal of investigative dermatology*. 131:1208-1215.

Fischer, J., A. Faure, B. Bouadjar, C. Blanchet-Bardon, A. Karaduman, I. Thomas, S. Emre, S. Cure, M. Ozguc, J. Weissenbach, and J.F. Prud'homme. 2000. Two new loci for autosomal recessive ichthyosis on chromosomes 3p21 and 19p12-q12 and evidence for further genetic heterogeneity. *American Journal of Human Genetics*. 66:904-913.

Flicek, P., M.R. Amode, D. Barrell, K. Beal, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fairley, S. Fitzgerald, L. Gil, L. Gordon, M. Hendrix, T. Hourlier, N. Johnson, A.K. Kahari, D. Keefe, S. Keenan, R. Kinsella, M. Komorowska, G. Koscielny, E. Kulesha, P. Larsson, I. Longden, W. McLaren, M. Muffato, B. Overduin, M. Pignatelli, B. Pritchard, H.S. Riat, G.R. Ritchie, M. Ruffier, M. Schuster, D. Sobral, Y.A. Tang, K. Taylor, S. Trevanion, J. Vandrovcova, S. White, M. Wilson, S.P. Wilder, B.L. Aken, E. Birney, F. Cunningham, I. Dunham, R. Durbin, X.M. Fernandez-Suarez, J. Harrow, J. Herrero, T.J. Hubbard, A. Parker, G. Proctor, G. Spudich, J. Vogel, A. Yates, A. Zadissa, and S.M. Searle. 2012. Ensembl 2012. *Nucleic Acids Research*. 40:D84-90.

Foitzik, K., G. Lindner, S. Mueller-Roever, M. Maurer, N. Botchkareva, V. Botchkarev, B. Handjiski, M. Metz, T. Hibino, T. Soma, G.P. Dotto, and R. Paus. 2000. Control of murine hair follicle regression (catagen) by TGF-beta1 in vivo. *Federation of American Societies for Experimental Biology Journal*. 14:752-760.

Folch, J., M. Lees, and G.H. Sloane Stanley. 1957. A simple method for the isolation and purification of total lipides from animal tissues. *The Journal of Biological Chemistry*. 226:497-509.

Freinkel, R.K., and T.N. Traczyk. 1983. Acid hydrolases of the epidermis: subcellular localization and relationship to cornification. *The Journal of Investigative Dermatology*. 80:441-446.

Freinkel, R.K., and T.N. Traczyk. 1985. Lipid composition and acid hydrolase content of lamellar granules of fetal rat epidermis. *The Journal of Investigative Dermatology*. 85:295-298.

Fuchs, E., and S. Raghavan. 2002. Getting under the skin of epidermal morphogenesis. *Nature Reviews: Genetics*. 3:199-209.

Fuchs-Telem, D., H. Stewart, D. Rapaport, J. Nousbeck, A. Gat, M. Gini, Y. Lugassy, S. Emmert, K. Eckl, H.C. Hennies, O. Sarig, D. Goldsher, B. Meilik, A. Ishida-

Yamamoto, M. Horowitz, and E. Sprecher. 2011. CEDNIK syndrome results from loss-of-function mutations in SNAP29. *The British Journal of Dermatology*. 164:610-616.

Fujimoto, W., K.W. Marvin, M.D. George, G. Celli, N. Darwiche, L.M. De Luca, and A.M. Jetten. 1993. Expression of cornifin in squamous differentiating epithelial tissues, including psoriatic and retinoic acid-treated skin. *The Journal of Investigative Dermatology*. 101:268-274.

Gallinaro, H., N. Jonca, L. Langbein, C. Vincent, M. Simon, G. Serre, and M. Guerrin. 2004. A 4.2 kb upstream region of the human corneodesmosin gene directs site-specific expression in hair follicles and hyperkeratotic epidermis of transgenic mice. *The Journal of Investigative Dermatology*. 122:730-738.

Garrod, D., and M. Chidgey. 2008. Desmosome structure, composition and function. *Biochimica et Biophysica Acta*. 1778:572-587.

Garrod, D., and T.E. Kimura. 2008. Hyper-adhesion: a new concept in cell-cell adhesion. *Biochemical Society transactions*. 36:195-201.

Gibbs, S., R. Fijneman, J. Wiegant, A.G. van Kessel, P. van De Putte, and C. Backendorf. 1993. Molecular characterization and evolution of the SPRR family of keratinocyte differentiation markers encoding small proline-rich proteins. *Genomics*. 16:630-637.

Gilissen, C., A. Hoischen, H.G. Brunner, and J.A. Veltman. 2012. Disease gene identification strategies for exome sequencing. *European journal of human genetics : EJHG*. 20:490-497.

Gish, W., and D.J. States. 1993. Identification of protein coding regions by database similarity search. *Nature Genetics*. 3:266-272.

Gissen, P., C.A. Johnson, N.V. Morgan, J.M. Stapelbroek, T. Forshew, W.N. Cooper, P.J. McKiernan, L.W. Klomp, A.A. Morris, J.E. Wraith, P. McClean, S.A. Lynch, R.J. Thompson, B. Lo, O.W. Quarrell, M. Di Rocco, R.C. Trembath, H. Mandel, S. Wali, F.E. Karet, A.S. Knisely, R.H. Houwen, D.A. Kelly, and E.R. Maher. 2004. Mutations in VPS33B, encoding a regulator of SNARE-dependent membrane fusion, cause arthrogryposis-renal dysfunction-cholestasis (ARC) syndrome. *Nature Genetics*. 36:400-404.

Godsel, L.M., R.P. Hobbs, and K.J. Green. 2008. Intermediate filament assembly: dynamics to disease. *Trends in cell biology*. 18:28-37.

Goldman, R.D., S. Khuon, Y.H. Chou, P. Opal, and P.M. Steinert. 1996. The function of intermediate filaments in cell shape and cytoskeletal integrity. *The Journal of Cell Biology*. 134:971-983.

Goldstein, D.B., A. Allen, J. Keebler, E.H. Margulies, S. Petrou, S. Petrovski, and S. Sunyaev. 2013. Sequencing studies in human genetics: design and interpretation. *Nature reviews. Genetics*. 14:460-470.

Goujon, M., H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, and R. Lopez. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic acids research*. 38:W695-699.

Grabowski, G.A. 2012. Gaucher disease and other storage disorders. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program*. 2012:13-18.

Grayson, S., A.G. Johnson-Winegar, B.U. Wintroub, R.R. Isseroff, E.H. Epstein, Jr., and P.M. Elias. 1985. Lamellar body-enriched fractions from neonatal mice: preparative techniques and partial characterization. *The Journal of Investigative Dermatology*. 85:289-294.

Gruber, R., P.M. Elias, D. Crumrine, T.-K. Lin, J.M. Brandner, J.-P. Hachem, R.B. Presland, P. Fleckman, A.R. Janecke, A. Sandilands, W.H.I. McLean, P.O. Fritsch,

M. Mildner, E. Tschachler, and M. Schmuth. 2011. Filaggrin Genotype in Ichthyosis Vulgaris Predicts Abnormalities in Epidermal Structure and Function. *The American Journal of Pathology*. 178:2252-2263.

Hachem, J.P., D. Crumrine, J. Fluhr, B.E. Brown, K.R. Feingold, and P.M. Elias. 2003. pH directly regulates epidermal permeability barrier homeostasis, and stratum corneum integrity/cohesion. *The Journal of Investigative Dermatology*. 121:345-353.

Haftek, M., S. Callejon, Y. Sandjeu, K. Padois, F. Falson, F. Pirot, P. Portes, F. Demarne, and V. Jannin. 2011. Compartmentalization of the human stratum corneum by persistent tight junction-like structures. *Experimental Dermatology*. 20:617-621.

Haines, R.L., and E.B. Lane. 2012. Keratins and disease at a glance. *Journal of cell science*. 125:3923-3928.

Halata, Z., M. Grim, and K.I. Bauman. 2003. Friedrich Sigmund Merkel and his "Merkel cell", morphology, development, and physiology: review and new results. *The Anatomical Record. Part A, Discoveries in molecular, cellular, and evolutionary biology*. 271:225-239.

Hamada, T., A. Sandilands, S. Fukuda, S. Sakaguchi, B. Ohyama, S. Yasumoto, W.H. McLean, and T. Hashimoto. 2008. De novo occurrence of the filaggrin mutation p.R501X with prevalent mutation c.3321delA in a Japanese family with ichthyosis vulgaris complicated by atopic dermatitis. *The Journal of Investigative Dermatology*. 128:1323-1325.

Hansen, L.A., N. Alexander, M.E. Hogan, J.P. Sundberg, A. Dlugosz, D.W. Threadgill, T. Magnuson, and S.H. Yuspa. 1997. Genetically null mice reveal a central role for epidermal growth factor receptor in the differentiation of the hair follicle and normal hair development. *The American Journal of Pathology*. 150:1959-1975.

Hart, P.S., Y. Zhang, E. Firatli, C. Uygur, M. Lotfazar, M.D. Michalec, J.J. Marks, X. Lu, B.J. Coates, W.K. Seow, R. Marshall, D. Williams, J.B. Reed, J.T. Wright, and T.C. Hart. 2000. Identification of cathepsin C mutations in ethnically diverse papillon-Lefevre syndrome patients. *Journal of Medical Genetics*. 37:927-932.

Hatzfeld, M., C. Haffner, K. Schulze, and U. Vinzens. 2000. The function of plakophilin 1 in desmosome assembly and actin filament organization. *The Journal of cell biology*. 149:209-222.

Hebert, J.M., T. Rosenquist, J. Gotz, and G.R. Martin. 1994. FGF5 as a regulator of the hair growth cycle: evidence from targeted and spontaneous mutations. *Cell*. 78:1017-1025.

Henry, J., E. Toulza, C.Y. Hsu, L. Pellerin, S. Balica, J. Mazereeuw-Hautier, C. Paul, G. Serre, N. Jonca, and M. Simon. 2012. Update on the epidermal differentiation complex. *Frontiers in bioscience*. 17:1517-1532.

Herrmann, T., F. van der Hoeven, H.J. Grone, A.F. Stewart, L. Langbein, I. Kaiser, G. Liebisch, I. Gosch, F. Buchkremer, W. Drobnik, G. Schmitz, and W. Stremmel. 2003. Mice with targeted disruption of the fatty acid transport protein 4 (Fatp 4, Slc27a4) gene show features of lethal restrictive dermopathy. *The Journal of Cell Biology*. 161:1105-1115.

Hitomi, K. 2005. Transglutaminases in skin epidermis. *European Journal of Dermatology*. 15:313-319.

Hobbs, R.P., and K.J. Green. 2012. Desmoplakin regulates desmosome hyperadhesion. *The Journal of investigative dermatology*. 132:482-485.

Hohl, D., P.A. de Viragh, F. Amiguet-Barras, S. Gibbs, C. Backendorf, and M. Huber. 1995. The small proline-rich proteins constitute a multigene family of

differentially regulated cornified cell envelope precursor proteins. *The Journal of Investigative Dermatology*. 104:902-909.

Holleran, W.M., E.I. Ginns, G.K. Menon, J.U. Grundmann, M. Fartasch, C.E. McKinney, P.M. Elias, and E. Sidransky. 1994. Consequences of beta-glucocerebrosidase deficiency in epidermis. Ultrastructure and permeability barrier alterations in Gaucher disease. *The Journal of Clinical Investigation*. 93:1756-1764.

Honari, S. 2004. Topical therapies and antimicrobials in the management of burn wounds. *Critical Care Nursing Clinics of North America*. 16:1-11.

Huber, M., I. Rettler, K. Bernasconi, E. Frenk, S.P. Lavrijsen, M. Ponec, A. Bon, S. Lautenschlager, D.F. Schorderet, and D. Hohl. 1995a. Mutations of keratinocyte transglutaminase in lamellar ichthyosis. *Science*. 267:525-528.

Huber, M., I. Rettler, K. Bernasconi, M. Wyss, and D. Hohl. 1995b. Lamellar ichthyosis is genetically heterogeneous--cases with normal keratinocyte transglutaminase. *The Journal of Investigative Dermatology*. 105:653-654.

Huizing, M., A. Didier, J. Walenta, Y. Anikster, W.A. Gahl, and H. Kramer. 2001. Molecular cloning and characterization of human VPS18, VPS 11, VPS16, and VPS33. *Gene*. 264:241-247.

Igarashi, S., T. Takizawa, Y. Yasuda, H. Uchiwa, S. Hayashi, H. Brysk, J.M. Robinson, K. Yamamoto, M.M. Brysk, and T. Horikoshi. 2004. Cathepsin D, but not cathepsin E, degrades desmosomes during epidermal desquamation. *The British Journal of Dermatology*. 151:355-361.

Igawa, S., M. Kishibe, M. Murakami, M. Honma, H. Takahashi, H. Iizuka, and A. Ishida-Yamamoto. 2011. Tight junctions in the stratum corneum explain spatial differences in corneodesmosome degradation. *Experimental Dermatology*. 20:53-57.

Ihrie, R.A., and L.D. Attardi. 2005. A new Perp in the lineup: linking p63 and desmosomal adhesion. *Cell Cycle*. 4:873-876.

Ishida-Yamamoto, A. 2003. Loricrin keratoderma: a novel disease entity characterized by nuclear accumulation of mutant loricrin. *Journal of dermatological science*. 31:3-8.

Ishida-Yamamoto, A., D. Hohl, D.R. Roop, H. Iizuka, and R.A. Eady. 1993. Loricrin immunoreactivity in human skin: localization to specific granules (L-granules) in acrosyringia. *Archives of Dermatological Research*. 285:491-498.

Ishida-Yamamoto, A., H. Kato, H. Kiyama, D.K. Armstrong, C.S. Munro, R.A. Eady, S. Nakamura, M. Kinouchi, H. Takahashi, and H. Iizuka. 2000. Mutant loricrin is not crosslinked into the cornified cell envelope but is translocated into the nucleus in loricrin keratoderma. *The Journal of investigative dermatology*. 115:1088-1094.

Ishida-Yamamoto, A., and M. Kishibe. 2011. Involvement of corneodesmosome degradation and lamellar granule transportation in the desquamation process. *Medical Molecular Morphology*. 44:1-6.

Ishida-Yamamoto, A., J.A. McGrath, H. Lam, H. Iizuka, R.A. Friedman, and A.M. Christiano. 1997. The molecular pathology of progressive symmetric erythrokeratoderma: a frameshift mutation in the loricrin gene and perturbations in the cornified cell envelope. *American Journal of Human Genetics*. 61:581-589.

Ishida-Yamamoto, A., M. Simon, M. Kishibe, Y. Miyauchi, H. Takahashi, S. Yoshida, T.J. O'Brien, G. Serre, and H. Iizuka. 2004. Epidermal lamellar granules transport different cargoes as distinct aggregates. *The Journal of Investigative Dermatology*. 122:1137-1144.

Israeli, S., H. Zamir, O. Sarig, R. Bergman, and E. Sprecher. 2011. Inflammatory peeling skin syndrome caused by a mutation in CDSN encoding corneodesmosin. *The Journal of Investigative Dermatology*. 131:779-781.

Ito, M. 1986. The innermost cell layer of the outer root sheath in anagen hair follicle: light and electron microscopic study. *Archives of Dermatological Research*. 279:112-119.

Iyengar, B. 2013. The melanocyte photosensory system in the human skin. *SpringerPlus*. 2:158.

Jang, S., T.H. Yang, E.J. An, H.K. Yoon, K.C. Sohn, A.Y. Cho, E.K. Ryu, Y.S. Park, T.Y. Yoon, J.H. Lee, and C.D. Kim. 2010. Role of plasminogen activator inhibitor-2 (PAI-2) in keratinocyte differentiation. *Journal of dermatological science*. 59:25-30.

Jansen, G.A., R. Ofman, S. Ferdinandusse, L. Ijlst, A.O. Muijsers, O.H. Skjeldal, O. Stokke, C. Jakobs, G.T. Besley, J.E. Wraith, and R.J. Wanders. 1997. Refsum disease is caused by mutations in the phytanoyl-CoA hydroxylase gene. *Nature Genetics*. 17:190-193.

Jarrett, A., and R.I. Spearman. 1957. The Keratin Defect and Hair-cycle of a New Mutant (Matted) in the House-mouse. *J. Embryol.exp.Morph*. 5:103-110.

Jarzab, J., B. Filipowska, J. Zebracka, M. Kowalska, A. Bozek, R. Rachowska, E. Gubala, A. Grzanka, E. Hadas, and B. Jarzab. 2010. Locus 1q21 Gene expression changes in atopic dermatitis skin lesions: deregulation of small proline-rich region 1A. *International archives of allergy and immunology*. 151:28-37.

Jobard, F., C. Lefevre, A. Karaduman, C. Blanchet-Bardon, S. Emre, J. Weissenbach, M. Ozguc, M. Lathrop, J.F. Prud'homme, and J. Fischer. 2002. Lipoxygenase-3 (ALOXE3) and 12(R)-lipoxygenase (ALOX12B) are mutated in non-bullous congenital ichthyosiform erythroderma (NCIE) linked to chromosome 17p13.1. *Human Molecular Genetics*. 11:107-113.

Jones, L.N., and D.E. Rivett. 1997. The role of 18-methyleicosanoic acid in the structure and formation of mammalian hair fibres. *Micron*. 28:469-485.

Jungersted, J.M., L.I. Hellgren, G.B. Jemec, and T. Agner. 2008. Lipids and skin barrier function--a clinical perspective. *Contact dermatitis*. 58:255-262.

Kalinin, A., L.N. Marekov, and P.M. Steinert. 2001. Assembly of the epidermal cornified cell envelope. *Journal of Cell Science*. 114:3069-3070.

Kalinin, A.E., W.W. Idler, L.N. Marekov, P. McPhie, B. Bowers, P.M. Steinert, and A.C. Steven. 2004. Co-assembly of envoplakin and periplakin into oligomers and Ca(2+)-dependent vesicle binding: implications for cornified cell envelope formation in stratified squamous epithelia. *The Journal of biological chemistry*. 279:22773-22780.

Kalinin, A.E., A.V. Kajava, and P.M. Steinert. 2002. Epithelial barrier function: assembly and structural features of the cornified cell envelope. *BioEssays*. 24:789-800.

Kam, E., K.A. Resing, S.K. Lim, and B.A. Dale. 1993. Identification of rat epidermal profilaggrin phosphatase as a member of the protein phosphatase 2A family. *Journal of Cell Science*. 106 ( Pt 1):219-226.

Kamatani, Y., K. Matsuda, Y. Okada, M. Kubo, N. Hosono, Y. Daigo, Y. Nakamura, and N. Kamatani. 2010. Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nature Genetics*. 42:210-215.

Kannan, K., L. Wang, J. Wang, M.M. Ittmann, W. Li, and L. Yen. 2011. Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. *Proceedings of the National Academy of Sciences of the United States of America*. 108:9172-9177.

Karashima, T., and F.M. Watt. 2002. Interaction of periplakin and envoplakin with intermediate filaments. *Journal of cell science*. 115:5027-5037.

Kartasova, T., N. Darwiche, Y. Kohno, H. Koizumi, S. Osada, N. Huh, U. Lichti, P.M. Steinert, and T. Kuroki. 1996. Sequence and expression patterns of mouse SPR1: Correlation of expression with epithelial function. *The Journal of investigative dermatology*. 106:294-304.

Kato, T., T. Takai, K. Mitsuishi, K. Okumura, and H. Ogawa. 2005. Cystatin A inhibits IL-8 production by keratinocytes stimulated with Der p 1 and Der f 1: biochemical skin barrier against mite cysteine proteases. *The Journal of allergy and clinical immunology*. 116:169-176.

Kauser, S., A.J. Thody, K.U. Schallreuter, C.L. Gummer, and D.J. Tobin. 2005. A fully functional proopiomelanocortin/melanocortin-1 receptor system regulates the differentiation of human scalp hair follicle melanocytes. *Endocrinology*. 146:532-543.

Kawasaki, H., K. Nagao, A. Kubo, T. Hata, A. Shimizu, H. Mizuno, T. Yamada, and M. Amagai. 2012. Altered stratum corneum barrier and enhanced percutaneous immune responses in filaggrin-null mice. *The Journal of Allergy and Clinical Immunology*. 129:1538-1546 e1536.

Kelsell, D.P., E.E. Norgett, H. Unsworth, M.T. Teh, T. Cullup, C.A. Mein, P.J. Dopping-Hepenstal, B.A. Dale, G. Tadini, P. Fleckman, K.G. Stephens, V.P. Sybert, S.B. Mallory, B.V. North, D.R. Witt, E. Sprecher, A.E. Taylor, A. Ilchyshyn, C.T. Kennedy, H. Goodyear, C. Moss, D. Paige, J.I. Harper, B.D. Young, I.M. Leigh, R.A. Eady, and E.A. O'Toole. 2005. Mutations in ABCA12 underlie the severe congenital skin disease harlequin ichthyosis. *American Journal of Human Genetics*. 76:794-803.

Kharfi, M., N. El Fekih, D. Ammar, H. Jaafoura, S. Schwonbeck, M.A. van Steensel, B. Fazaa, M.R. Kamoun, and J. Fischer. 2009. A missense mutation in TGM5 causes acral peeling skin syndrome in a Tunisian family. *The Journal of Investigative Dermatology*. 129:2512-2515.

Khnykin, D., J.H. Miner, and F. Jahnsen. 2011. Role of fatty acid transporters in epidermis: Implications for health and disease. *Dermato-endocrinology*. 3:53-61.

Kirschner, N., and J.M. Brandner. 2012. Barriers and more: functions of tight junction proteins in the skin. *Annals of the New York Academy of Sciences*. 1257:158-166.

Kirschner, N., C. Poetzl, P. von den Driesch, E. Wladykowski, I. Moll, M.J. Behne, and J.M. Brandner. 2009. Alteration of tight junction proteins is an early event in psoriasis: putative involvement of proinflammatory cytokines. *The American Journal of Pathology*. 175:1095-1106.

Kitajima, Y. 2013. Regulation and impairments of dynamic desmosome and corneodesmosome remodeling. *European journal of dermatology : EJD*.

Koch, P.J., P.A. de Viragh, E. Scharer, D. Bundman, M.A. Longley, J. Bickenbach, Y. Kawachi, Y. Suga, Z. Zhou, M. Huber, D. Hohl, T. Kartasova, M. Jarnik, A.C. Steven, and D.R. Roop. 2000. Lessons from loricrin-deficient mice: compensatory mechanisms maintaining skin barrier function in the absence of a major cornified envelope protein. *The Journal of cell biology*. 151:389-400.

Korge, B.P., A. Ishida-Yamamoto, C. Punter, P.J. Dopping-Hepenstal, H. Iizuka, A. Stephenson, R.A. Eady, and C.S. Munro. 1997. Loricrin mutation in Vohwinkel's keratoderma is unique to the variant with ichthyosis. *The Journal of Investigative Dermatology*. 109:604-610.

Krebsova, A., W. Kuster, G.G. Lestringant, B. Schulze, B. Hinz, P.M. Frossard, A. Reis, and H.C. Hennies. 2001. Identification, by homozygosity mapping, of a novel locus for autosomal recessive congenital ichthyosis on chromosome 17p, and evidence for further genetic heterogeneity. *American Journal of Human Genetics*. 69:216-222.

Kubo, A., K. Nagao, M. Yokouchi, H. Sasaki, and M. Amagai. 2009. External antigen uptake by Langerhans cells with reorganization of epidermal tight junction barriers. *The Journal of Experimental Medicine*. 206:2937-2946.

Landmann, L. 1986. Epidermal permeability barrier: transformation of lamellar granule-disks into intercellular sheets by a membrane-fusion process, a freeze-fracture study. *The Journal of Investigative Dermatology*. 87:202-209.

Lane, P.W. 1972. Two New Mutations in Linkage Group XVI of the House Mouse: Flaky tail and varitint-waddler-J. *Journal of Heredity*. 63:135-140.

Lane, P.W., and E.M. Eicher. 1979. Gene order in linkage group XVI of the house mouse. *Journal of Heredity*. 70:239-244.

Langbein, L., and J. Schweizer. 2005. Keratins of the human hair follicle. *International Review of Cytology*. 243:1-78.

Langbein, L., H. Yoshida, S. Praetzel-Wunder, D.A. Parry, and J. Schweizer. 2010. The keratins of the human beard hair medulla: the riddle in the middle. *The Journal of Investigative Dermatology*. 130:55-73.

Langlois, S., A.C. Maher, J.L. Manias, Q. Shao, G.M. Kidder, and D.W. Laird. 2007. Connexin levels regulate keratinocyte differentiation in the epidermis. *The Journal of biological chemistry*. 282:30171-30180.

Larabell, C.A., K. Fukuyama, and W.L. Epstein. 1993. Desmosome differentiation during epidermal cornification: new observations obtained from intermediate voltage electron microscopy. *The Journal of Investigative Dermatology*. 101:103-104.

Lechler, T., and E. Fuchs. 2005. Asymmetric cell divisions promote stratification and differentiation of mammalian skin. *Nature*. 437:275-280.

Lechler, T., and E. Fuchs. 2007. Desmoplakin: an unexpected regulator of microtubule organization in the epidermis. *The Journal of Cell Biology*. 176:147-154.

Lee, J., and T. Tumbar. 2012. Hairy tale of signaling in hair follicle development and cycling. *Seminars in cell & developmental biology*. 23:906-916.

Lee, W.S. 2011. Integral hair lipid in human hair follicle. *Journal of dermatological science*. 64:153-158.

Lefevre, C., S. Audebert, F. Jobard, B. Bouadjar, H. Lakhdar, O. Boughdene-Stambouli, C. Blanchet-Bardon, R. Heilig, M. Foglio, J. Weissenbach, M. Lathrop, J.F. Prud'homme, and J. Fischer. 2003. Mutations in the transporter ABCA12 are associated with lamellar ichthyosis type 2. *Human Molecular Genetics*. 12:2369-2378.

Lefevre, C., F. Jobard, F. Caux, B. Bouadjar, A. Karaduman, R. Heilig, H. Lakhdar, A. Wollenberg, J.L. Verret, J. Weissenbach, M. Ozguc, M. Lathrop, J.F. Prud'homme, and J. Fischer. 2001. Mutations in CGI-58, the gene encoding a new protein of the esterase/lipase/thioesterase subfamily, in Chanarin-Dorfman syndrome. *American Journal of Human Genetics*. 69:1002-1012.

Levy-Nissenbaum, E., R.C. Betz, M. Frydman, M. Simon, H. Lahat, T. Bakhan, B. Goldman, A. Bygum, M. Pierick, A.M. Hillmer, N. Jonca, J. Toribio, R. Kruse, G. Dewald, S. Cichon, C. Kubisch, M. Guerrin, G. Serre, M.M. Nothen, and E. Pras. 2003. Hypotrichosis simplex of the scalp is associated with nonsense mutations in CDSN encoding corneodesmosin. *Nature Genetics*. 34:151-153.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 25:1754-1760.

Li, S., K. Nikulina, J. DeVoss, A.J. Wu, E.C. Strauss, M.S. Anderson, and N.A. McNamara. 2008. Small proline-rich protein 1B (SPRR1B) is a biomarker for squamous metaplasia in dry eye disease. *Investigative ophthalmology & visual science*. 49:34-41.

Lindner, G., V.A. Botchkarev, N.V. Botchkareva, G. Ling, C. van der Veen, and R. Paus. 1997. Analysis of apoptosis during hair follicle regression (catagen). *The American Journal of Pathology*. 151:1601-1617.

Lobo, I. 2008. Same Genetic Mutation, Different Genetic Disease Phenotype. *Nature Education*. 1.

Ma, L., J. Liu, T. Wu, M. Plikus, T.X. Jiang, Q. Bi, Y.H. Liu, S. Muller-Rover, H. Peters, J.P. Sundberg, R. Maxson, R.L. Maas, and C.M. Chuong. 2003. 'Cyclic alopecia' in Msx2 mutants: defects in hair cycling and hair shaft differentiation. *Development*. 130:379-389.

Maatta, A., T. DiColandrea, K. Groot, and F.M. Watt. 2001. Gene targeting of envoplakin, a cytoskeletal linker protein and precursor of the epidermal cornified envelope. *Molecular and cellular biology*. 21:7047-7053.

Maestrini, E., A.P. Monaco, J.A. McGrath, A. Ishida-Yamamoto, C. Camisa, A. Hovnanian, D.E. Weeks, M. Lathrop, J. Uitto, and A.M. Christiano. 1996. A molecular defect in loricrin, the major component of the cornified cell envelope, underlies Vohwinkel's syndrome. *Nature Genetics*. 13:70-77.

Mandal, T.K., and D.T. Downing. 1993. Freeze-fracture electron microscopic and osmotic water permeability studies of epidermal lipid liposomes derived from stratum corneum lipids of porcine epidermis. *Acta Dermato-Venereologica*. 73:12-17.

Marekov, L.N., and P.M. Steinert. 1998. Ceramides are bound to structural proteins of the human foreskin epidermal cornified cell envelope. *The Journal of Biological Chemistry*. 273:17763-17770.

Marenholz, I., M. Zirra, D.F. Fischer, C. Backendorf, A. Ziegler, and D. Mischke. 2001. Identification of human epidermal differentiation complex (EDC)-encoded genes by subtractive hybridization of entire YACs to a gridded keratinocyte cDNA library. *Genome Research*. 11:341-355.

Marshall, D., M.J. Hardman, K.M. Nield, and C. Byrne. 2001. Differentially expressed late constituents of the epidermal cornified envelope. *Proceedings of the National Academy of Sciences of the United States of America*. 98:13031-13036.

Matsumoto, K., M. Muto, S. Seki, T. Saida, N. Horiuchi, H. Takahashi, A. Ishida-Yamamoto, and H. Iizuka. 2001. Loricrin keratoderma: a cause of congenital ichthyosiform erythroderma and collodion baby. *The British journal of dermatology*. 145:657-660.

McAleer, M.A., and A.D. Irvine. 2013. The multifunctional role of filaggrin in allergic skin disease. *The Journal of allergy and clinical immunology*. 131:280-291.

McLean, W.H., S.M. Morley, E.B. Lane, R.A. Eady, W.A. Griffiths, D.G. Paige, J.I. Harper, C. Higgins, and I.M. Leigh. 1994. Ichthyosis bullosa of Siemens--a disease involving keratin 2e. *The Journal of Investigative Dermatology*. 103:277-281.

McLean, W.H., E.L. Rugg, D.P. Lunny, S.M. Morley, E.B. Lane, O. Swensson, P.J. Dopping-Hepenstal, W.A. Griffiths, R.A. Eady, C. Higgins, and et al. 1995. Keratin 16 and keratin 17 mutations cause pachyonychia congenita. *Nature Genetics*. 9:273-278.

Mehrel, T., D. Hohl, J.A. Rothnagel, M.A. Longley, D. Bundman, C. Cheng, U. Lichti, M.E. Bisher, A.C. Steven, P.M. Steinert, and et al. 1990. Identification of a major keratinocyte cell envelope protein, loricrin. *Cell*. 61:1103-1112.

Meisler, M.H., and M.F. Seldin. 1991. Mouse chromosome 3. *Mammalian Genome : Official Journal of the International Mammalian Genome Society*. 1 Spec No:S42-50.

Menon, G.K., K.R. Feingold, and P.M. Elias. 1992. Lamellar body secretory response to barrier disruption. *The Journal of Investigative Dermatology*. 98:279-289.

Menon, G.K., L.F. Price, B. Bommannan, P.M. Elias, and K.R. Feingold. 1994. Selective obliteration of the epidermal calcium gradient leads to enhanced lamellar body secretion. *The Journal of investigative dermatology*. 102:789-795.

Meyer, L.R., A.S. Zweig, A.S. Hinrichs, D. Karolchik, R.M. Kuhn, M. Wong, C.A. Sloan, K.R. Rosenbloom, G. Roe, B. Rhead, B.J. Raney, A. Pohl, V.S. Malladi, C.H. Li, B.T. Lee, K. Learned, V. Kirkup, F. Hsu, S. Heitner, R.A. Harte, M. Haeussler, L. Guruvadoo, M. Goldman, B.M. Giardine, P.A. Fujita, T.R. Dreszer, M. Diekhans, M.S. Cline, H. Clawson, G.P. Barber, D. Haussler, and W.J. Kent. 2013. The UCSC Genome Browser database: extensions and updates 2013. *Nucleic acids research*. 41:D64-69.

Mihalik, S.J., J.C. Morrell, D. Kim, K.A. Sacksteder, P.A. Watkins, and S.J. Gould. 1997. Identification of PAHX, a Refsum disease gene. *Nature Genetics*. 17:185-189.

Mils, V., C. Vincent, F. Croute, and G. Serre. 1992. The expression of desmosomal and corneodesmosomal antigens shows specific variations during the terminal differentiation of epidermis and hair follicle epithelia. *The Journal of Histochemistry and Cytochemistry : Official Journal of the Histochemistry Society*. 40:1329-1337.

Mizutani, Y., S. Mitsutake, K. Tsuji, A. Kihara, and Y. Igarashi. 2009. Ceramide biosynthesis in keratinocyte and its role in skin function. *Biochimie*. 91:784-790.

Mobraaten, L.E., H.P. Bunker, J. DeMaeyer-Gulgnard, E. DeMaeyer, and D.W. Balley. 1984. Location of histocompatibility and interferon loci on chromosome 3 of the mouse. *Journal of Heredity*. 75:233-234.

Molin, S., S. Vollmer, E.H. Weiss, P. Weisenseel, T. Ruzicka, and J.C. Prinz. 2011. Deletion of the late cornified envelope genes LCE3B and LCE3C may promote chronic hand eczema with allergic contact dermatitis. *Journal of investigational allergology & clinical immunology*. 21:472-479.

Moniaga, C.S., G. Egawa, H. Kawasaki, M. Hara-Chikuma, T. Honda, H. Tanizaki, S. Nakajima, A. Otsuka, H. Matsuoka, A. Kubo, J. Sakabe, Y. Tokura, Y. Miyachi, M. Amagai, and K. Kabashima. 2010. Flaky tail mouse denotes human atopic dermatitis in the steady state and by topical application with Dermatophagoides pteronyssinus extract. *The American Journal of Pathology*. 176:2385-2393.

Moniaga, C.S., S.K. Jeong, G. Egawa, S. Nakajima, M. Hara-Chikuma, J.E. Jeon, S.H. Lee, T. Hibino, Y. Miyachi, and K. Kabashima. 2013. Protease activity enhances production of thymic stromal lymphopoietin and basophil accumulation in flaky tail mice. *The American Journal of Pathology*. 182:841-851.

Morita, K., M.E. Hogan, L.B. Nanney, L.E. King, Jr., M. Manabe, T.T. Sun, and J.P. Sundberg. 1995. Cutaneous ultrastructural features of the flaky skin (fsn) mouse mutation. *The Journal of Dermatology*. 22:385-395.

Mueller, A., J. O'Rourke, J. Grimm, K. Guillemin, M.F. Dixon, A. Lee, and S. Falkow. 2003. Distinct gene expression profiles characterize the histopathological stages of disease in Helicobacter-induced mucosa-associated lymphoid tissue

lymphoma. *Proceedings of the National Academy of Sciences of the United States of America*. 100:1292-1297.

Nemes, Z., M. Demeny, L.N. Marekov, L. Fesus, and P.M. Steinert. 2000. Cholesterol 3-sulfate interferes with cornified envelope assembly by diverting transglutaminase 1 activity from the formation of cross-links and esters to the hydrolysis of glutamine. *The Journal of Biological Chemistry*. 275:2636-2646.

Nemes, Z., L.N. Marekov, L. Fesus, and P.M. Steinert. 1999a. A novel function for transglutaminase 1: attachment of long-chain omega-hydroxyceramides to involucrin by ester bond formation. *Proceedings of the National Academy of Sciences of the United States of America*. 96:8402-8407.

Nemes, Z., L.N. Marekov, and P.M. Steinert. 1999b. Involucrin cross-linking by transglutaminase 1. Binding to membranes directs residue specificity. *The Journal of Biological Chemistry*. 274:11013-11021.

Nemes, Z., and P.M. Steinert. 1999. Bricks and mortar of the epidermal barrier. *Experimental & Molecular Medicine*. 31:5-19.

Nicol, J.W., G.A. Helt, S.G. Blanchard, Jr., A. Raja, and A.E. Loraine. 2009. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics*. 25:2730-2731.

Niessen, C.M. 2007. Tight junctions/adherens junctions: basic structure and function. *The Journal of Investigative Dermatology*. 127:2525-2532.

Nishifuji, K., and J.S. Yoon. 2013. The stratum corneum: the rampart of the mammalian body. *Veterinary dermatology*. 24:60-72 e15-66.

Nomura, T., A. Sandilands, M. Akiyama, H. Liao, A.T. Evans, K. Sakai, M. Ota, H. Sugiura, K. Yamamoto, H. Sato, C.N. Palmer, F.J. Smith, W.H. McLean, and H. Shimizu. 2007. Unique mutations in the filaggrin gene in Japanese patients with ichthyosis vulgaris and atopic dermatitis. *The Journal of Allergy and Clinical Immunology*. 119:434-440.

Nusrat, A., C. von Eichel-Streiber, J.R. Turner, P. Verkade, J.L. Madara, and C.A. Parkos. 2001. Clostridium difficile toxins disrupt epithelial barrier function by altering membrane microdomain localization of tight junction proteins. *Infection and Immunity*. 69:1329-1336.

Oji, V., K.M. Eckl, K. Aufenvenne, M. Natebus, T. Tarinski, K. Ackermann, N. Seller, D. Metze, G. Nurnberg, R. Folster-Holst, M. Schafer-Korting, I. Hausser, H. Traupe, and H.C. Hennies. 2010a. Loss of corneodesmosin leads to severe skin barrier defect, pruritus, and atopy: unraveling the peeling skin disease. *American Journal of Human Genetics*. 87:274-281.

Oji, V., G. Tadini, M. Akiyama, C. Blanchet Bardon, C. Bodemer, E. Bourrat, P. Coudiere, J.J. DiGiovanna, P. Elias, J. Fischer, P. Fleckman, M. Gina, J. Harper, T. Hashimoto, I. Hausser, H.C. Hennies, D. Hohl, A. Hovnanian, A. Ishida-Yamamoto, W.K. Jacyk, S. Leachman, I. Leigh, J. Mazereeuw-Hautier, L. Milstone, F. Morice-Picard, A.S. Paller, G. Richard, M. Schmuth, H. Shimizu, E. Sprecher, M. Van Steensel, A. Taieb, J.R. Toro, P. Vabres, A. Vahlquist, M. Williams, and H. Traupe. 2010b. Revised nomenclature and classification of inherited ichthyoses: results of the First Ichthyosis Consensus Conference in Soreze 2009. *Journal of the American Academy of Dermatology*. 63:607-641.

Orwin, D.F. 1971. Cell differentiation in the lower outer sheath of the Romney wool follicle: a companion cell layer. *Australian Journal of Biological Sciences*. 24:989-999.

Orwin, D.F. 1979. The cytology and cytochemistry of the wool follicle. *International Review of Cytology*. 60:331-374.

Orwin, D.F., R.W. Thomson, and N.E. Flower. 1973. Plasma membrane differentiations of keratinizing cells of the wool follicle. II. Desmosomes. *Journal of Ultrastructure Research*. 45:15-29.

Osorio, F., M. Leao, F. Azevedo, and S. Magina. 2012. Lamellar Ichthyosis Due to ALOX12B Mutation. *Actas Dermo-Sifiliograficas*.

Oyoshi, M.K., G.F. Murphy, and R.S. Geha. 2009. Filaggrin-deficient mice exhibit TH17-dominated skin inflammation and permissiveness to epicutaneous sensitization with protein antigen. *The Journal of Allergy and Clinical Immunology*. 124:485-493, 493 e481.

Palmer, C.N., A.D. Irvine, A. Terron-Kwiatkowski, Y. Zhao, H. Liao, S.P. Lee, D.R. Goudie, A. Sandilands, L.E. Campbell, F.J. Smith, G.M. O'Regan, R.M. Watson, J.E. Cecil, S.J. Bale, J.G. Compton, J.J. DiGiovanna, P. Fleckman, S. Lewis-Jones, G. Arseculeratne, A. Sergeant, C.S. Munro, B. El Houate, K. McElreavey, L.B. Halkjaer, H. Bisgaard, S. Mukhopadhyay, and W.H. McLean. 2006. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nature Genetics*. 38:441-446.

Parmentier, L., C. Blanchet-Bardon, S. Nguyen, J.F. Prud'homme, L. Dubertret, and J. Weissenbach. 1995. Autosomal recessive lamellar ichthyosis: identification of a new mutation in transglutaminase 1 and evidence for genetic heterogeneity. *Human Molecular Genetics*. 4:1391-1395.

Parmentier, L., H. Lakhdar, C. Blanchet-Bardon, S. Marchand, L. Dubertret, and J. Weissenbach. 1996. Mapping of a second locus for lamellar ichthyosis to chromosome 2q33-35. *Human Molecular Genetics*. 5:555-559.

Petiot, A., F.J. Conti, R. Grose, J.M. Revest, K.M. Hodivala-Dilke, and C. Dickson. 2003. A crucial role for Fgfr2-IIIb signalling in epidermal development and hair follicle patterning. *Development*. 130:5493-5501.

Petit, E., M. Huber, A. Rochat, C. Bodemer, D. Teillac-Hamel, J.P. Muh, J. Revuz, Y. Barrandon, M. Lathrop, Y. de Prost, D. Hohl, and A. Hovnanian. 1997. Three novel point mutations in the keratinocyte transglutaminase (TGK) gene in lamellar ichthyosis: significance for mutant transcript level, TGK immunodetection and activity. *European Journal of Human Genetics*. 5:218-228.

Pigors, M., D. Kiritsi, C. Cobzaru, A. Schwieger-Briel, J. Suarez, F. Faletra, H. Aho, L. Makela, J.S. Kern, L. Bruckner-Tuderman, and C. Has. 2012. TGM5 mutations impact epidermal differentiation in acral peeling skin syndrome. *The Journal of Investigative Dermatology*. 132:2422-2429.

Porter, R.M., and E.B. Lane. 2003. Phenotypes, genotypes and their contribution to understanding keratin function. *Trends In Genetics*. 19:278-285.

Powell, B.C., and J.S. Beltrame. 1994. Characterization of a hair (wool) keratin intermediate filament gene domain. *The Journal of Investigative Dermatology*. 102:171-177.

Powell, B.C., and G.E. Rogers. 1997. The role of keratin proteins and their genes in the growth, structure and properties of hair. *EXS*. 78:59-148.

Presland, R.B., D. AU - Boggess, S.P. AU - Lewis, C. AU - Hull, P. AU - Fleckman, and J.P. AU - Sundberg. 2000. Loss of Normal Profilaggrin and Filaggrin in Flaky Tail (ft//ft) Mice: an Animal Model for the Filaggrin-Deficient Skin Disease Ichthyosis Vulgaris. *Journal of Investigative Dermatology*. 115.

Proksch, E., K.R. Feingold, and P.M. Elias. 1992. Epidermal HMG CoA reductase activity in essential fatty acid deficiency: barrier requirements rather than eicosanoid generation regulate cholesterol synthesis. *The Journal of Investigative Dermatology*. 99:216-220.

Proksch, E., B. Johanna M, and J.-M. Jensen. 2008. The skin: an indispensable barrier. *Experimental Dermatology*. 17:1063-1072.

Raghunath, M., H.C. Hennies, B. Ahvazi, M. Vogel, A. Reis, P.M. Steinert, and H. Traupe. 2003. Self-healing collodion baby: a dynamic phenotype explained by a particular transglutaminase-1 mutation. *The Journal of investigative dermatology*. 120:224-228.

Reis, A., H.C. Hennies, L. Langbein, M. Digweed, D. Mischke, M. Drechsler, E. Schrock, B. Royer-Pokora, W.W. Franke, K. Sperling, and et al. 1994. Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK). *Nature Genetics*. 6:174-179.

Resing, K.A., R.S. Johnson, and K.A. Walsh. 1993. Characterization of protease processing sites during conversion of rat profilaggrin to filaggrin. *Biochemistry*. 32:10036-10045.

Reynolds, A.J., and C.A. Jahoda. 1991. Hair follicle stem cells? A distinct germinative epidermal cell population is activated in vitro by the presence of hair dermal papilla cells. *Journal of Cell Science*. 99 ( Pt 2):373-385.

Rice, R.H., V.J. Wong, and K.E. Pinkerton. 1994. Ultrastructural visualization of cross-linked protein features in epidermal appendages. *Journal of Cell Science*. 107 ( Pt 7):1985-1992.

Rizzo, W.B. 2013. Fatty aldehyde and fatty alcohol metabolism: Review and importance for epidermal structure and function. *Biochimica et biophysica acta*.

Robinson, N.A., S. Lapic, J.F. Welter, and R.L. Eckert. 1997. S100A11, S100A10, annexin I, desmosomal proteins, small proline-rich proteins, plasminogen activator inhibitor-2, and involucrin are components of the cornified envelope of cultured human epidermal keratinocytes. *The Journal of Biological Chemistry*. 272:12035-12046.

Rogers, G.E. 1959. Electron microscope studies of hair and wool. *Annals of the New York Academy of Sciences*. 83:378-399.

Rogers, G.E. 2004. Hair follicle differentiation and regulation. *The International Journal of Developmental Biology*. 48:163-170.

Romero, S., G. Grompone, N. Carayol, J. Mounier, S. Guadagnini, M.C. Prevost, P.J. Sansonetti, and G.T. Van Nhieu. 2011. ATP-mediated Erk1/2 activation stimulates bacterial capture by filopodia, which precedes Shigella invasion of epithelial cells. *Cell Host & Microbe*. 9:508-519.

Rothnagel, J.A., A.M. Dominey, L.D. Dempsey, M.A. Longley, D.A. Greenhalgh, T.A. Gagne, M. Huber, E. Frenk, D. Hohl, and D.R. Roop. 1992. Mutations in the rod domains of keratins 1 and 10 in epidermolytic hyperkeratosis. *Science*. 257:1128-1130.

Ruhrberg, C., M.A. Hajibagheri, D.A. Parry, and F.M. Watt. 1997. Periplakin, a novel component of cornified envelopes and desmosomes that belongs to the plakin family and forms complexes with envoplakin. *The Journal of Cell Biology*. 139:1835-1849.

Sajic, D., R. Asiniwasis, and S. Skotnicki-Grant. 2012. A look at epidermal barrier function in atopic dermatitis: physiologic lipid replacement and the role of ceramides. *Skin therapy letter*. 17:6-9.

Sakai, K., M. Akiyama, Y. Sugiyama-Nakagiri, J.R. McMillan, D. Sawamura, and H. Shimizu. 2007. Localization of ABCA12 from Golgi apparatus to lamellar granules in human upper epidermal keratinocytes. *Experimental Dermatology*. 16:920-926.

Salomon, D., E. Masgrau, S. Vischer, S. Ullrich, E. Dupont, P. Sappino, J.H. Saurat, and P. Meda. 1994. Topography of mammalian connexins in human skin. *The Journal of Investigative Dermatology*. 103:240-247.

Sandilands, A., G.M. O'Regan, H. Liao, Y. Zhao, A. Terron-Kwiatkowski, R.M. Watson, A.J. Cassidy, D.R. Goudie, F.J. Smith, W.H. McLean, and A.D. Irvine. 2006. Prevalent and rare mutations in the gene encoding filaggrin cause ichthyosis vulgaris and predispose individuals to atopic dermatitis. *The Journal of Investigative Dermatology*. 126:1770-1775.

Sandilands, A., C. Sutherland, A.D. Irvine, and W.H. McLean. 2009. Filaggrin in the frontline: role in skin barrier function and disease. *Journal of Cell Science*. 122:1285-1294.

Sandilands, A., A. Terron-Kwiatkowski, P.R. Hull, G.M. O'Regan, T.H. Clayton, R.M. Watson, T. Carrick, A.T. Evans, H. Liao, Y. Zhao, L.E. Campbell, M. Schmuth, R. Gruber, A.R. Janecke, P.M. Elias, M.A. van Steensel, I. Nagtzaam, M. van Geel, P.M. Steijlen, C.S. Munro, D.G. Bradley, C.N. Palmer, F.J. Smith, W.H. McLean, and A.D. Irvine. 2007. Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nature Genetics*. 39:650-654.

Sando, G.N., H. Zhu, J.M. Weis, J.T. Richman, P.W. Wertz, and K.C. Madison. 2003. Caveolin expression and localization in human keratinocytes suggest a role in lamellar granule biogenesis. *The Journal of Investigative Dermatology*. 120:531-541.

Sasaki, T., A. Shiohama, A. Kubo, H. Kawasaki, A. Ishida-Yamamoto, T. Yamada, T. Hachiya, A. Shimizu, H. Okano, J. Kudoh, and M. Amagai. 2013. A homozygous nonsense mutation in the gene for Tmem79, a component for the lamellar granule secretory system, produces spontaneous eczema in an experimental model of atopic dermatitis. *The Journal of allergy and clinical immunology*.

Scharschmidt, T.C., M.Q. Man, Y. Hatano, D. Crumrine, R. Gunathilake, J.P. Sundberg, K.A. Silva, T.M. Mauro, M. Hupe, S. Cho, Y. Wu, A. Celli, M. Schmuth, K.R. Feingold, and P.M. Elias. 2009. Filaggrin deficiency confers a paracellular barrier abnormality that reduces inflammatory thresholds to irritants and haptens. *The Journal of Allergy and Clinical Immunology*. 124:496-506, 506 e491-496.

Schmuth, M., G. Yosipovitch, M.L. Williams, F. Weber, H. Hintner, S. Ortiz-Urda, K. Rappersberger, D. Crumrine, K.R. Feingold, and P.M. Elias. 2001. Pathogenesis of the permeability barrier abnormality in epidermolytic hyperkeratosis. *Journal of Investigative Dermatology*. 117:837-847.

Searle, A.G., and R.I. Spearman. 1957. 'Matted', a New Hair-mutant in the House-mouse: Genetics and Morphology. *J. Embryol. exp. Morph*:93-102.

Senshu, T., S. Kan, H. Ogawa, M. Manabe, and H. Asaga. 1996. Preferential deimination of keratin K1 and filaggrin during the terminal differentiation of human epidermis. *Biochemical and Biophysical Research Communications*. 225:712-719.

Serre, G., V. Mils, M. Haftek, C. Vincent, F. Croute, A. Reano, J.P. Ouhayoun, S. Bettinger, and J.P. Soleilhavoup. 1991. Identification of late differentiation antigens of human cornified epithelia, expressed in re-organized desmosomes and bound to cross-linked envelope. *The Journal of Investigative Dermatology*. 97:1061-1072.

Sevilla, L.M., R. Nachat, K.R. Groot, J.F. Klement, J. Uitto, P. Djian, A. Maatta, and F.M. Watt. 2007. Mice deficient in involucrin, envoplakin, and periplakin have a defective epidermal barrier. *The Journal of cell biology*. 179:1599-1612.

Shwayder, T., S. Conn, and L. Lowe. 1997. Acral peeling skin syndrome. *Archives of Dermatology*. 133:535-536.

Sievers, F., A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J.D. Thompson, and D.G. Higgins. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*. 7:539.

Simon, M., N. Jonca, M. Guerrin, M. Haftek, D. Bernard, C. Caubet, T. Egelrud, R. Schmidt, and G. Serre. 2001. Refined characterization of corneodesmosin proteolysis during terminal differentiation of human epidermis and its relationship to desquamation. *The Journal of Biological Chemistry*. 276:20292-20299.

Simpson, C.L., D.M. Patel, and K.J. Green. 2011. Deconstructing the skin: cytoarchitectural determinants of epidermal morphogenesis. *Nature reviews. Molecular cell biology*. 12:565-580.

Skjeldal, O.H., O. Stokke, S. Refsum, J. Norseth, and H. Petit. 1987. Clinical and biochemical heterogeneity in conditions with phytanic acid accumulation. *Journal of the Neurological Sciences*. 77:87-96.

Slominski, A., D.J. Tobin, S. Shibahara, and J. Wortsman. 2004. Melanin pigmentation in mammalian skin and its hormonal regulation. *Physiological Reviews*. 84:1155-1228.

Smih, F., F. Desmoulin, M. Berry, A. Turkieh, R. Harmancey, J. Iacovoni, C. Trouillet, C. Delmas, A. Pathak, O. Lairez, F. Koukoui, P. Massabuau, J. Ferrieres, M. Galinier, and P. Rouet. 2011. Blood signature of pre-heart failure: a microarrays study. *PloS one*. 6:e20414.

Smith, F.J., A.D. Irvine, A. Terron-Kwiatkowski, A. Sandilands, L.E. Campbell, Y. Zhao, H. Liao, A.T. Evans, D.R. Goudie, S. Lewis-Jones, G. Arseculeratne, C.S. Munro, A. Sergeant, G. O'Regan, S.J. Bale, J.G. Compton, J.J. DiGiovanna, R.B. Presland, P. Fleckman, and W.H. McLean. 2006. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nature Genetics*. 38:337-342.

Smith, F.J., M.F. Jonkman, H. van Goor, C.M. Coleman, S.P. Covello, J. Uitto, and W.H. McLean. 1998. A mutation in human keratin K6b produces a phenocopy of the K17 disorder pachyonychia congenita type 2. *Human Molecular Genetics*. 7:1143-1148.

Smith, K.R., and D.M. Thiboutot. 2008. Thematic review series: skin lipids. Sebaceous gland lipids: friend or foe? *Journal of Lipid Research*. 49:271-281.

Sprecher, E., S. Chavanas, J.J. DiGiovanna, S. Amin, K. Nielsen, J.S. Prendiville, R. Silverman, N.B. Esterly, M.K. Spraker, E. Guelig, M.L. de Luna, M.L. Williams, B. Buehler, E.C. Siegfried, L. Van Maldergem, E. Pfendner, S.J. Bale, J. Uitto, A. Hovnanian, and G. Richard. 2001. The spectrum of pathogenic mutations in SPINK5 in 19 families with Netherton syndrome: implications for mutation detection and first case of prenatal diagnosis. *The Journal of Investigative Dermatology*. 117:179-187.

Sprecher, E., A. Ishida-Yamamoto, M. Mizrahi-Koren, D. Rapaport, D. Goldsher, M. Indelman, O. Topaz, I. Chefetz, H. Keren, J. O'Brien T, D. Bercovich, S. Shalev, D. Geiger, R. Bergman, M. Horowitz, and H. Mandel. 2005. A mutation in SNAP29, coding for a SNARE protein involved in intracellular trafficking, causes a novel neurocutaneous syndrome characterized by cerebral dysgenesis, neuropathy, ichthyosis, and palmoplantar keratoderma. *American Journal of Human Genetics*. 77:242-251.

Steinert, P.M., E. Candi, T. Kartasova, and L. Marekov. 1998a. Small proline-rich proteins are cross-bridging proteins in the cornified cell envelopes of stratified squamous epithelia. *Journal of Structural Biology*. 122:76-85.

Steinert, P.M., E. Candi, E. Tarcsa, L.N. Marekov, M. Sette, M. Paci, B. Ciani, P. Guerrieri, and G. Melino. 1999. Transglutaminase crosslinking and structural studies of the human small proline rich 3 protein. *Cell Death and Differentiation*. 6:916-930.

Steinert, P.M., S.I. Chung, and S.Y. Kim. 1996a. Inactive zymogen and highly active proteolytically processed membrane-bound forms of the transglutaminase 1 enzyme in human epidermal keratinocytes. *Biochemical and Biophysical Research Communications*. 221:101-106.

Steinert, P.M., T. Kartasova, and L.N. Marekov. 1998b. Biochemical evidence that small proline-rich proteins and trichohyalin function in epithelia by modulation of the biomechanical properties of their cornified cell envelopes. *The Journal of Biological Chemistry*. 273:11758-11769.

Steinert, P.M., S.Y. Kim, S.I. Chung, and L.N. Marekov. 1996b. The transglutaminase 1 enzyme is variably acylated by myristate and palmitate during differentiation in epidermal keratinocytes. *The Journal of Biological Chemistry*. 271:26242-26250.

Steinert, P.M., J.W. Mack, B.P. Korge, S.Q. Gan, S.R. Haynes, and A.C. Steven. 1991. Glycine loops in proteins: their occurrence in certain intermediate filament chains, loricrins and single-stranded RNA binding proteins. *International Journal of Biological Macromolecules*. 13:130-139.

Steinert, P.M., and L.N. Marekov. 1995. The proteins elafin, filaggrin, keratin intermediate filaments, loricrin, and small proline-rich proteins 1 and 2 are isodipeptide cross-linked components of the human epidermal cornified cell envelope. *The Journal of Biological Chemistry*. 270:17702-17711.

Steingrimsson, E., N.G. Copeland, and N.A. Jenkins. 2005. Melanocyte stem cell maintenance and hair graying. *Cell*. 121:9-12.

Steven, A.C., M.E. Bisher, D.R. Roop, and P.M. Steinert. 1990. Biosynthetic pathways of filaggrin and loricrin--two major proteins expressed by terminally differentiated epidermal keratinocytes. *Journal of Structural Biology*. 104:150-162.

Stone, D.L., N. Tayebi, E. Orvisky, B. Stubblefield, V. Madike, and E. Sidransky. 2000. Glucocerebrosidase gene mutations in patients with type 2 Gaucher disease. *Human Mutation*. 15:181-188.

Stout, C.E., J.L. Costantin, C.C. Naus, and A.C. Charles. 2002. Intercellular calcium signaling in astrocytes via ATP release through connexin hemichannels. *The Journal of biological chemistry*. 277:10482-10488.

Suga, Y., M. Jarnik, P.S. Attar, M.A. Longley, D. Bundman, A.C. Steven, P.J. Koch, and D.R. Roop. 2000. Transgenic mice expressing a mutant form of loricrin reveal the molecular basis of the skin diseases, Vohwinkel syndrome and progressive symmetric erythrokeratoderma. *The Journal of Cell Biology*. 151:401-412.

Sun, L.D., F.L. Xiao, Y. Li, W.M. Zhou, H.Y. Tang, X.F. Tang, H. Zhang, H. Schaarschmidt, X.B. Zuo, R. Foelster-Holst, S.M. He, M. Shi, Q. Liu, Y.M. Lv, X.L. Chen, K.J. Zhu, Y.F. Guo, D.Y. Hu, M. Li, Y.H. Zhang, X. Zhang, J.P. Tang, B.R. Guo, H. Wang, Y. Liu, X.Y. Zou, F.S. Zhou, X.Y. Liu, G. Chen, L. Ma, S.M. Zhang, A.P. Jiang, X.D. Zheng, X.H. Gao, P. Li, C.X. Tu, X.Y. Yin, X.P. Han, Y.Q. Ren, S.P. Song, Z.Y. Lu, X.L. Zhang, Y. Cui, J. Chang, M. Gao, X.Y. Luo, P.G. Wang, X. Dai, W. Su, H. Li, C.P. Shen, S.X. Liu, X.B. Feng, C.J. Yang, G.S. Lin, Z.X. Wang, J.Q. Huang, X. Fan, Y. Wang, Y.X. Bao, S. Yang, J.J. Liu, A. Franke, S. Weidinger, Z.R. Yao, and X.J. Zhang.

2011. Genome-wide association study identifies two new susceptibility loci for atopic dermatitis in the Chinese Han population. *Nature genetics*. 43:690-694.

Szeverenyi, I., A.J. Cassidy, C.W. Chung, B.T. Lee, J.E. Common, S.C. Ogg, H. Chen, S.Y. Sim, W.L. Goh, K.W. Ng, J.A. Simpson, L.L. Chee, G.H. Eng, B. Li, D.P. Lunny, D. Chuon, A. Venkatesh, K.H. Khoo, W.H. McLean, Y.P. Lim, and E.B. Lane. 2008. The Human Intermediate Filament Database: comprehensive information on a gene family involved in many human diseases. *Human Mutation*. 29:351-360.

Tarcsa, E., E. Candi, T. Kartasova, W.W. Idler, L.N. Marekov, and P.M. Steinert. 1998. Structural and transglutaminase substrate properties of the small proline-rich 2 family of cornified cell envelope proteins. *The Journal of Biological Chemistry*. 273:23297-23303.

Tarcsa, E., L.N. Marekov, G. Mei, G. Melino, S.C. Lee, and P.M. Steinert. 1996. Protein unfolding by peptidylarginine deiminase. Substrate specificity and structural relationships of the natural substrates trichohyalin and filaggrin. *The Journal of Biological Chemistry*. 271:30709-30716.

Terrinoni, A., P. Puddu, B. Didona, V. De Laurenzi, E. Candi, F.J. Smith, W.H. McLean, and G. Melino. 2000. A mutation in the V1 domain of K16 is responsible for unilateral palmoplantar verrucous nevus. *The Journal of Investigative Dermatology*. 114:1136-1140.

Thauvin-Robinet, C., F. Mugneret, P. Callier, M. Chouchane, E. Garron, E. Manceau, N. Laurent, C. Durand, A. Nivelon-Chevallier, E. Sapin, F. Huet, and L. Faivre. 2005. Unique survival in chrondrodysplasia-hermaphrodism syndrome. *American Journal of Medical Genetics*. 132A:335-337.

Thiboutot, D. 2004. Regulation of human sebaceous glands. *The Journal of Investigative Dermatology*. 123:1-12.

Thomas, A.C., T. Cullup, E.E. Norgett, T. Hill, S. Barton, B.A. Dale, E. Sprecher, E. Sheridan, A.E. Taylor, R.S. Wilroy, C. DeLozier, N. Burrows, H. Goodyear, P. Fleckman, K.G. Stephens, L. Mehta, R.M. Watson, R. Graham, R. Wolf, A. Slavotinek, M. Martin, D. Bourn, C.A. Mein, E.A. O'Toole, and D.P. Kelsell. 2006. ABCA12 is the major harlequin ichthyosis gene. *The Journal of Investigative Dermatology*. 126:2408-2413.

Toomes, C., J. James, A.J. Wood, C.L. Wu, D. McCormick, N. Lench, C. Hewitt, L. Moynihan, E. Roberts, C.G. Woods, A. Markham, M. Wong, R. Widmer, K.A. Ghaffar, M. Pemberton, I.R. Hussein, S.A. Temtamy, R. Davies, A.P. Read, P. Sloan, M.J. Dixon, and N.S. Thakker. 1999. Loss-of-function mutations in the cathepsin C gene result in periodontal disease and palmoplantar keratosis. *Nature Genetics*. 23:421-424.

Tsuji, S., P.V. Choudary, B.M. Martin, B.K. Stubblefield, J.A. Mayor, J.A. Barranger, and E.I. Ginns. 1987. A mutation in the human glucocerebrosidase gene in neuronopathic Gaucher's disease. *The New England Journal of Medicine*. 316:570-575.

Tu, C.L., W. Chang, and D.D. Bikle. 2001. The extracellular calcium-sensing receptor is required for calcium-induced differentiation in human keratinocytes. *The Journal of biological chemistry*. 276:41079-41085.

Tybulewicz, V.L., M.L. Tremblay, M.E. LaMarca, R. Willemsen, B.K. Stubblefield, S. Winfield, B. Zablocka, E. Sidransky, B.M. Martin, S.P. Huang, and et al. 1992. Animal model of Gaucher's disease from targeted disruption of the mouse glucocerebrosidase gene. *Nature*. 357:407-410.

Uitto, J., G. Richard, and J.A. McGrath. 2007. Diseases of epidermal keratins and their linker proteins. *Experimental cell research*. 313:1995-2009.

Villadangos, J.A., and P. Schnorrer. 2007. Intrinsic and cooperative antigen-presenting functions of dendritic-cell subsets in vivo. *Nature Reviews: Immunology*. 7:543-555.

Wang, H., D.A. Parry, L.N. Jones, W.W. Idler, L.N. Marekov, and P.M. Steinert. 2000. In vitro assembly and structure of trichocyte keratin intermediate filaments: a novel role for stabilization by disulfide bonding. *The Journal of cell biology*. 151:1459-1468.

Wang, S., and Y. Wang. 2013. Peptidylarginine deiminases in citrullination, gene regulation, health and pathogenesis. *Biochimica et biophysica acta*. 1829:1126-1135.

Waterham, H.R., and J.M. Cregg. 1997. Peroxisome biogenesis. *BioEssays*. 19:57-66.

Wells, R.S., and C.B. Kerr. 1966. Clinical features of autosomal dominant and sex-linked ichthyosis in an English population. *British Medical Journal*. 1:947-950.

Wertz, P.W. 1997. Integral lipids of hair and stratum corneum. *EXS*. 78:227-237.

Wigderson, M., N. Firon, Z. Horowitz, S. Wilder, Y. Frishberg, O. Reiner, and M. Horowitz. 1989. Characterization of mutations in Gaucher patients by cDNA cloning. *American Journal of Human Genetics*. 44:365-377.

Wilson, N.J., A.G. Messenger, S.A. Leachman, E.A. O'Toole, E.B. Lane, W.H. McLean, and F.J. Smith. 2010. Keratin K6c mutations cause focal palmoplantar keratoderma. *The Journal of Investigative Dermatology*. 130:425-429.

Winter, H., L. Langbein, S. Praetzel, M. Jacobs, M.A. Rogers, I.M. Leigh, N. Tidman, and J. Schweizer. 1998. A novel human type II cytokeratin, K6hf, specifically expressed in the companion layer of the hair follicle. *The Journal of Investigative Dermatology*. 111:955-962.

Xu, J., and B.J. Nicholson. 2013. The role of connexins in ear and skin physiology - functional insights from disease-associated mutations. *Biochimica et biophysica acta*. 1828:167-178.

Yamaguchi, T., N. Omatsu, S. Matsushita, and T. Osumi. 2004. CGI-58 interacts with perilipin and is localized to lipid droplets. Possible involvement of CGI-58 mislocalization in Chanarin-Dorfman syndrome. *The Journal of Biological Chemistry*. 279:30490-30497.

Yanagi, T., M. Akiyama, H. Nishihara, K. Sakai, W. Nishie, S. Tanaka, and H. Shimizu. 2008. Harlequin ichthyosis model mouse reveals alveolar collapse and severe fetal skin barrier defects. *Human molecular genetics*. 17:3075-3083.

Yoneda, K., T. Nakagawa, O.T. Lawrence, J. Huard, T. Demitsu, Y. Kubota, and R.B. Presland. 2012. Interaction of the profilaggrin N-terminal domain with loricrin in human cultured keratinocytes and epidermis. *The Journal of investigative dermatology*. 132:1206-1214.

Zhang, X. 2012. Genome-wide association study of skin complex diseases. *Journal of dermatological science*. 66:89-97.

Zouboulis, C.C., W.C. Chen, M.J. Thornton, K. Qin, and R. Rosenfield. 2007. Sexual hormones in human skin. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme*. 39:85-95.

# APPENDIX I

## Immune profile of the *Matt*$^{ma/ma}$ mouse

# Appendix I

## Immune profile of *Matt*$^{ma/ma}$ (*matted*) mice (Performed by Padraic Fallon et. al., Trinity College, Dublin)

These *Matt*$^{ma/ma}$ mice show diffuse epidermal hyperplasia (P <0.001) and hyperkeratosis of anuclear granular cells (orthokeratosis) (P < 0.001) (Figure 1.g) with dermal inflammatory infiltrates (P <0.001) (Figure 1.d). A more severe phenotype compared to *Flg*$^{ft/ft}$ mice which only had occasional foci of acanthosis, mild diffuse orthokeratosis and dermal infiltrates, and presented with no overt dermatitis (Figure 1.g). Trans-epidermal water loss (TEWL) is a method of measurement used to analyzed severity of skin inflammation of skin barrier integrity in atopic dermatitis patients (Eberlein-Konig et al., 2000; Gupta et al., 2008). In the *Matt*$^{ma/ma}$ and DM mice, TEWL values were significantly raised indicating a dysregulation in skin barrier integrity, but were unaltered in *Flg*$^{ft/ft}$ mice (Figure 1.f). Spontaneous elevated IgE levels were seen in all mutants, but were exaggerated in *Matt*$^{ma/ma}$ and DM animals. House dust mite allergen was applied as a model to demonstrate percutaneous sensitization in *Matt*$^{ma/ma}$ as compared to wild type mice, where there was an increase in IgE, IgG1 and IgG2a responses following HDM allergen application in *Matt*$^{ma/ma}$, that was not seen in *Flg*$^{ft/ft}$ and wild type mice, indicating that *Matt*$^{ma/ma}$ mice are more susceptible to percutaneous allergic sensitization due to a defective skin barrier (Figure 1.a-c). As shown in previous studies (Jarrett and Spearman, 1957; Searle and Spearman, 1957), the *Matt*$^{ma/ma}$ and DM mice demonstrated what appears to be a keratinizing defect, causing clumping of hairs with brittles and longitudinal splitting. They also had distorted hair follicle morphogenesis with defective cuticle morphology.

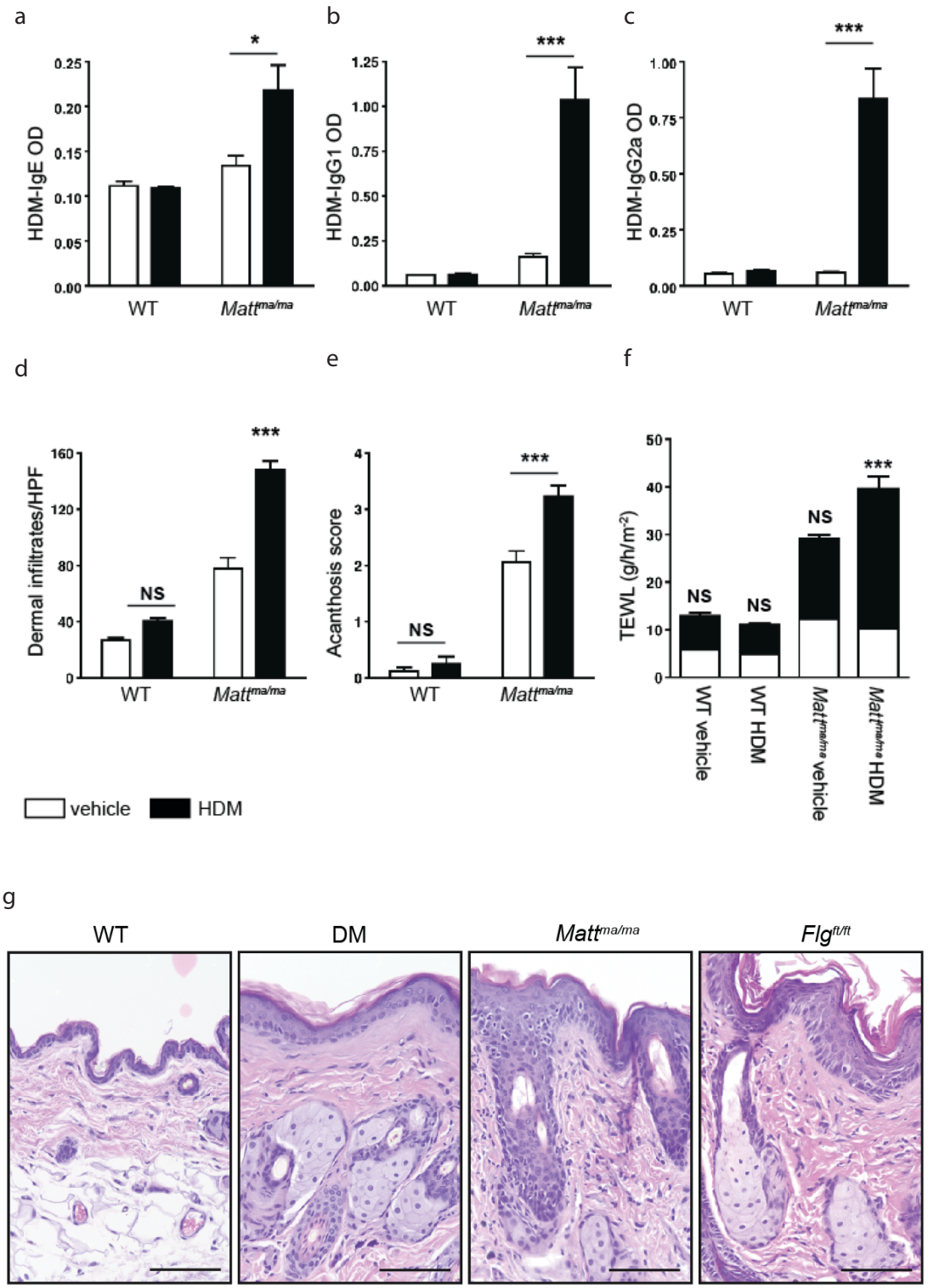



**Figure I** Immune profile of matted mouse showing elevated levels of IgE, IgG1, IgG2a (a - c), following house dust mite (HDM) challenge. Comparison of dermal infiltrates and acanthosis levels (d, e). Transepidermal water loss was markedly elevated in *ma/ma* mouse when challenged with both HDM and vehicle. H&E sections showing gross hyperplasia and hyperkeratosis in DM and *ma/ma* (g).

# APPENDIX II

## Complete list of mapping primers designed to cover murine chromosome 3: 87,000,000 – 95,000,000

## Appendix II

Complete list of mapping primers designed to cover murine chromosome 3: 87,000,000 – 95,000,000.

| Gene/mRNA | | Primer (5' - 3') | Chromosomal location |
|---|---|---|---|
| gm16000 | exon 1 | CAA AAA ACA GAT TAG GGG TCA G | chr3:81040296-81040658 |
| | | CAT TGT TAT TCC ATA CCC CG | |
| | exon 2 | ATG GGG TTC AGA GGC TAA G | chr3:81037803-81038391 |
| | | GGC TAT GGA AAG GAT TGG | |
| | | CAA TCC TTT CCA TAG CCC | chr3:81037290-81037819 |
| | | TTA TGT GGA AAC TAC CCT GC | |
| gm9989 | exon 1 | CAC ACT CCA TCA TTG TCT CC | chr3:81924017-81924625 |
| | | CAG CAA GTG TCT TTA CCA GC | |
| | exon 2 | ATG GAC CCT GTG ATA GGT GGT G | chr3:81923414-81924163 |
| | | TGT GTG TCC ATC CAT CCA GC | |
| | exon 3 | GGA ATC TTA CTG TAT TGG GCT C | chr3:81922892-81923479 |
| | | TCA ACT CTT CAG AAA CCC G | |
| | exon 4 | AAG AGT TGA TGC CAC TTC C | chr3:81922357-81922900 |
| | | TGT GGA TGC CAA GTA AGC | |
| 4930564K09rik | exon 1 | GAA AGG GAC TTG GAG TGG | chr3:82876609+82877254 |
| | | TAG TGT ATT AGA TGT GGC TCA G | |
| | exon 2 | CAT AGA ACA GGA TAG TAG CAC AG | chr3:82878581+82879201 |
| | | CAG TGC AAA GCT GTG GTT | |
| | exon 3 | GAT GCA TGC ATA CTC CTA AGG | chr3:82940566+82943774 |
| | | GCA TCT TCC ACT TCA AGG G | |
| | | GCA GAC CTA TTG CCG TTT | |
| | | GAT CCT AGT TCC TTC TCC TCA C | |
| | | CAG ATA TAC ATG GGA TTT AAG G | |
| | | TGG ATG AGA CTG CCA TTC | |
| | | CAG CTA TAC TAT ATG TCC TGA A | |
| | | GGA TGA TGA GGG CAT AGA A | |
| | | CTG TGA CTC TAG TTC CAA GGG | |
| | | CAG TCA GCA GGT TGT TAA GG | |
| | | GCT ACC TGG AAC ACA ATG A | |
| | | TAG GAC CAT ATC CAT TAG AC | |
| gm10710 | | CCA CTT TCC AAG GAC AAC T | chr3:83126696-83129323 |
| | | CTT AGA CTT GGT GCT GCT G | |
| | | GCA GCA GCA CCA AGT CTA AG | |
| | | CGA ATG TCA CCC ACC AGT GT | |
| | | TCA TCC GAG TCC TCC GAA AG | |
| | | CAT CGC TTT TTC TCT CTT CTT C | |
| | | GCA AAC AAA CGA AAA GAC C | |
| | | GCC GTT TTT TAG TCC TTC TAC | |
| | | GTA GAA GGA CTA AAA AAC GGC | |

| | | | |
|---|---|---|---|
| | | GGT CAA TCC GCA CAT AAA TC | |
| gm16790 | | GAG AGA AGG AGG AAC GAG AAA | chr3:83594006-83790084 |
| | | CAC CGT TCT CCC TCT CCC TTT | |
| | | CCT GGC TTC CAG TGG ATG TTC | |
| | | AAG CTA CAA TGA TTT AGT TTG | |
| | | CCA GGT CAA CTA CTA CGT ATA T | |
| | | TGT TGG CTT ATA CCG TGC AC | |
| D930015E06Rik | | GCT GGG CTA CGA TTC ATT C | chr3:84039893-84040547 |
| | | GCG AAC CCT CAC AGA TTT C | |
| | | GGC TGG GCA TTG TAT TTG | chr3:84038966-84039558 |
| | | TTA GAG ACA GCC GTG GAG AC | |
| | | CCA GAG GGT AGA GTT ACA GGT G | chr3:84031374-84031902 |
| | | TCT GAA CAA GAG GGA AGG C | |
| | | CTA TTG ATG ATG GAG ACT TTG G | chr3:83962026-83962661 |
| | | CCA AAC TTA GGG CTT CTC TG | |
| | | GAG AAG CCC TAA GTT TGG TC | chr3:83961429-83962043 |
| | | AAA GGG CAG GAA ATC TCT C | |
| | | GGT CCC ACA TCA GAA GAG AAG | chr3:83960282-83960941 |
| | | TGA GCC CAT AGC CCA ATA G | |
| | | GCT GTT TGG AAC TCA CTG TAT G | chr3:83942312-83942920 |
| | | TTA TGG ATG AGG CAG GAC C | |
| | | TGG AGA GTC TAT GGC TTT TTG | chr3:83940435-83941001 |
| | | CCT GAG TGA AAC TGC CAA C | |
| | | CCT CCC CAA ACT CTT CTT TAC | chr3:83938618-83939116 |
| | | AAC TGG GGT TCA GGA GAA G | |
| | | GGT ATT TCC TGA GTT GTA TTG G | chr3:83937853-83938395 |
| | | GCC ACA ACT GAA CTA AAA GC | |
| | | TTT AGT CAT AGC GAG TGC G | chr3:83937180-83937645 |
| | | GCA CAG AGG AAG GAA ACT G | |
| | | CAC TGA TGA TTG TCC AGA GG | chr3:83935737-83936288 |
| | | TCT CTC ACA AAA GCA CAA GC | |
| | | CTC ATA GTT TGT GAA GTT TGG C | chr3:83934432-83935031 |
| | | TGC TGG CTG AAG TGC TAA C | |
| | | CAT TCC TCC TCT CCT CTG AC | chr3:83931539-83931982 |
| | | GCA CAG TGA CAA AGA AGG TG | |
| | | TGT AGG ATG AGG TCA AGA GTT G | chr3:83930288-83930819 |
| | | CAC TAA AAA GGA ATG CTA AGG G | |
| | | TGG CAT TTT AGT CTG TGA GAG | chr3:83926853-83927293 |
| | | GGG AAA GAG TCA AAC AAC TTC | |
| | | AAG TGA AGT GTA TGA AAC CTG C | chr3:83897985-83898545 |
| | | CCA TAA GGT CCT CTT CTA AAG C | |
| gm6525 | | AAT AGT AGC CAT AAA GGT CGC | chr3:84174453+84175072 |
| | | TTT CTC TTC TTG TCG CCT C | |
| BM937755 | | GCT TGG ACT CTA AAC AAA CCA C | chr3:84257587+84258259 |
| | | AAG AGC AGG AGT GAT TCA GAG | |
| | | ACA GGA CGG GAA TGT AGT G | chr3:84258158+84258764 |
| | | AGG GCA CAG ATA GCG TTA G | |

| | | | |
|---|---|---|---|
| | | CAA ACT GTC ACA AGA ACA ACC | chr3:84258516+84259070 |
| | | CTC TTT GAA GGA ACT ACT CAG C | |
| | | AAA GGC AAC ATC TCC ACC | chr3:84258960+84259573 |
| | | GGG ACT GAG CAC AGA ATG | |
| BC048678 | exon 1 | GAA ACA GAC ATT TAG GCC TAG T | chr3:85317409+85317943 |
| | | TCT AAT GGA AAG CGA GCC | |
| | exon 2 | GTA TTT AAA TAC CCA AGG TG | chr3:85330853+85331456 |
| | | CTA TAG ACT CTG CCA TGC T | |
| | exon 3 | GGC TTT GGT ATG TTA ACC AT | chr3:85331687+85332127 |
| | | AGT TAG AAG TGT CCC AGG C | |
| AK086005 | exon 1 | GAG ACA ACA GAC AGC TTC | chr3:85553065-85553635 |
| | | TCC CTG CAT ATA CAT ATG | |
| | exon 2 | CAG GTA TCC AGA AGC ACT GG | chr3:85544182-85544765 |
| | | CTT TTC TAA AAG GAA GGG TTT G | |
| gm15535 | exon 2 | GTG GTT AAG CCC CTT CTA C | chr3:85556976-85557580 |
| | | TGA TTA TAC AGT GTT CTG CC | |
| | | CAT GGC AGC ATA TGA GTG TG | chr3:85556576-85557170 |
| | | CAC GGC CTC TGC ATC AG | |
| | | CTC AAG CAG GTC AGG AAG CAG | chr3:85556066-85556616 |
| | | AAC CTG GAG TTC ACT GAT TTG G | |
| | | AGA CAG ATC CCT GGG ATT ACA G | chr3:85555776-85556122 |
| | | TTG CTT TTG AAT CAC ATG GC | |
| gm9790 | | AGT CAA ACC ATT GGC TCT C | chr3:85915526+85916108 |
| | | TCT AAC TAA AGC AAG CAC TTC C | |
| gm3788 | | AAA TCA GTG GAC CAG AAG G | chr3:86671655-86672325 |
| | | CAG CAT TTT CTC ATT GTG C | |
| BY319657 | | CCC TCT CAT TCT GAC CAC TTC | chr3:87365512+87366166 |
| | | ATG ACC ATC TGG GGC ATA G | |
| BE847318 | | GCC CTC TTC TTA TGT TTG C | chr3:87382566+87383140 |
| | | GCT CGT GTA TGA ACT ACC CTA C | |
| AW494178 | | CTT CAT TTC CAG TCA TAC ACG | chr3:87445869+87446464 |
| | | TAA TAC TCC CCT GCG TCA C | |
| DY243455 | | CCC TCC ACA TTT GTT CTA TTC | chr3:87531735+87532371 |
| | | CAA GGT TTC TCT TTT AGC CC | |
| CK137060 | | CCT TCT TTC TAC GAG CAC G | chr3:87539251+87539704 |
| | | CCT CAC CAC TAC TGT TCA CAC | |
| CF805084 | | ATC TCT CCA GCC CTG AAA C | chr3:87557673+87558299 |
| | | CGG ACC TAA AAC CTT CAC AC | |
| BE865208 | | AGA GAG ACT TTC GGA GCC TG | chr3:87617775+87618380 |
| | | GCA GAC AGA TAA AAG CCA TCC | |
| gm6570 | | GGG AGG AAG CAA AGA AGA G | chr3:87679693-87680329 |
| | | AGA ATC ACC TGG GTC AAT G | |
| AK040918 | | GTG TGT CTG GAG ATG GCT AC | chr3:87691123+87691806 |
| | | TGA GAA AGC AGG GTT GAC | |
| | | TTC TGC TTC ATA GCC ACC G | chr3:87692246+87692842 |
| | | CGG TTC CTG TAA GCA GTC AC | |
| | | AAA CAG TCT CAC GGT CAC C | chr3:87692772+ |

| | | | |
|---|---|---|---|
| | | GGG TAT GTA TGA ACA GGT GC | 87693383 |
| | | GCT GGA GAG ATG ATT CTA TGG | chr3:87693223+ 87693805 |
| | | TAC TTC CAG GTT TCT CTC CTC | |
| 4933430H15rik | exon 1 | GCG AAA TGG TAA AGA TGG G | chr3:87748307-87748942 |
| | | GGA TTT TTT GGC TCG TCC | |
| | exon 2-3 | GCA AAG CAG CTT CCA CTT G | chr3:87745537-87746118 |
| | | AGG CCA AAG TCA TCT CAG G | |
| | exon 4-5 | CTC ACG GGC TCA CAA TTC A | chr3:87745004-87745592 |
| | | TCC CTG TGA TTC CTA CAG AGA | |
| | exon 6-7 | CTA TCC TTA GGC TCA GAT | chr3:87742874-87743469 |
| | | AGA ACA GAG TGA CTT GTA AG | |
| | exon 8 | GCT TAC AAG TCA CTC TGT TCT | chr3:87742256-87742894 |
| | | CAG CAG ATT TTG TTG AAC TA | |
| | exon 9 | CAA AGT GAG CCA TTT CCC TGT | chr3:87741495-87742044 |
| | | CCA GAG CCC AGC TTC TCT T | |
| | exon 10 | CAT TAT AGC TGT CCC CAT GT | chr3:87740600-87741197 |
| | | AGG AAG ACC TCC TAA GGA AG | |
| | exon 11 | CAT CTG ACA TCA CTG GCA GCC | chr3:87739494+ 87739747 |
| | | GGG ATG GAC TTG ATT CTC CAA | |
| | exon 12 | GTT TCT CCA GCA CCT GGG AC | chr3:87739252-87740073 |
| | | GTT TTA GGG CTG TTG AAC GG | |
| | exon 13 | GTT TGT TTC CTG CAA TTA G | chr3:87738724-87739323 |
| | | CAA ACA CAT GGA GAT AGC | |
| | | GGA ACC CAG CCT TCT CTG A | chr3:87738174-87738761 |
| | | CCG GTG TGT AGA CCT TCT CAG | |
| | | CTG GGC CAG AGC CAG TTG T | chr3:87737594-87738223 |
| | | GGG AGG TCT CCC AGA TTC CA | |
| | | GAC ACT GAG AAG GCC TAG G | chr3:87737003-87737668 |
| | | TCA AAT GTA GCC AGC GTC | |
| | | CTG TGG GAA GAT GGT GAG G | chr3:87736640-87737043 |
| | | CCC CCA GAA CCT TCT ATT TAT G | |
| AK185634 | | GGA AAC CAA TAG TGA GTG AGC | chr3:87749141+ 87749692 |
| | | TCT TTG CTC AGA GGA TTG C | |
| BM940369 | | TTC GGG ACA GGA AGG TTA C | chr3:87778640+ 87779197 |
| | | CAA GCA CTG AGA GAA GGA ATG | |
| BB365020 | | CGT TTT GTG TTC AAC CTG AC | chr3:87804416+ 87804935 |
| | | AAC CTC TGT TGC CCT CTT C | |
| BB160298 | | TTA CCC ACT CTC TCC AAG G | chr3:87849217+ 87849865 |
| | | TTT TGA AAG TGG AGG TTC C | |
| AA008399 | | CCC ACA TCA ACC ATC AAG | chr3:87899438+ 87900008 |
| | | CAA AAC CCA GGT CCT TAT G | |
| CJ043528 | | GGA CTT CCA GTT GCT CTC TG | chr3:87911129+ 87911777 |
| | | CCT AAG CAA TGT GAA GGC TG | |
| BY607175 | | TCT GGT CAA GGG ACT TCT AAC | chr3:87922507+ 87923081 |
| | | CCC ATC CTC AAC TTA CCA G | |
| S78708 | | AAG GGG TGG TGG TGG TGT TTA C | chr3:87970546+ 87971229 |
| | | AGA TTC TTC CCC GAC GCA AC | |

| | | | |
|---|---|---|---|
| EL608440 | | TTC AGA CAC AGC GAT GAG C | chr3:87992458+87992926 |
| | | TTT TCT GTC TCT CCC CGT G | |
| | | GAC GGA AGT GAC AGA ATG G | chr3:87992869+87993475 |
| | | TCA GAA ACA GAG CAT TCA CTC | |
| CN720784 | | AGT TCC AGG ATA GCC AAT G | chr3:88026129+88026681 |
| | | TAC CGC AAA GTG CCT AAG | |
| | | GGA CTC CAT TCA CAC AAG C | chr3:88026551+88027108 |
| | | TGA CTT AGT GTG CTG CGT G | |
| CF583689 | | GGG GTC TTG CTA CAA TGA TAC | chr3:88056339+88056956 |
| | | GTG CTC AAA CCT GGC TAA G | |
| BB200528 | | GCC TGA TTA TGC TAT GTA GAC C | chr3:88069327+88069991 |
| | | GAA CTT CCC GAT AGA CAC TTG | |
| BG807888 | | GGT AGT TCA GTC GGT CCT TC | chr3:88107777+88108347 |
| | | TAG GGG CAA ACC AAT GAG | |
| | | ACC TAT GAA AAC AGA GGC G | chr3:88108203+88108817 |
| | | ATG GGC TGA GTC TGT GAA C | |
| 1700113A16rik | exon 1 | CGG AAA TGG ATG TTA GGC | chr3:88177289-88177907 |
| | | TCC AGC CTT GTG ATT GTG | |
| | exon 2 | AGG TGG ATT TCT GCG TTC | chr3:88175660-88176184 |
| | | GGT GTA GTG TAG CCT GGT TG | |
| | exon 3 | TTG CTG CTT CGG CAC CAA | chr3:88172102-88172505 |
| | | TGA TTT CTT TGG TGC GGC G | |
| | | AGC CTG CCC TCC AAA TCA | chr3:88171578-88172180 |
| | | GTT TGG AGA GAC AGA AAT GAT G | |
| | | CCC CTT CTG CTT ATG TAG AGG | chr3:88171435-88171732 |
| | | TCA GGG ATA GTG GGA CCG | |
| 1700021C14Rik | exon 1 | GAA AGG CAA GGC AAG TTA GAT AG | chr3:88296353-88297031 |
| | | AGA GCA ATA GAT GGT CCC G | |
| | exon 2 | CCA GAC ACC TCT ATG AGA TAG | chr3:88295020-88295572 |
| | | CCA AGT AAC CAA GTA GCA G | |
| | exon 3 | CCT GCT GGA TAC ACT CAA TCA | chr3:88282674-88283282 |
| | | CAC ACA CAC AGA ACA TAA CGC | |
| 0610031J06rik | exon 1 | AAA TAC CTG TTC ACC TTG AGC | chr3:88324912+88325432 |
| | | TGT TCC TGG GCT ACA ATG | |
| | exon 2 | GCC TGT GAG CAT ACC TGA | chr3:88325624+88326072 |
| | | TTT CTC TTG AAT ATC CCT CC | |
| | exon 3-4 | CAA GCC TCT TCC ACT CTA CC | chr3:88326085+88326668 |
| | | TTA CAT TCC CAC CCA TCT ATC | |
| | exon 5 | CCT TTT AGT ATG GGG TTG C | chr3:88327511+88328106 |
| | | GGT AGT ATT GGT CCC AGT AGC | |
| | exon 6 | GGC TAC TGG GAC CAA TAC TAC | chr3:88328085+88328717 |
| | | CCA CAC TGC ACA TCC ATA C | |
| tmem79 | exon 1 | TTT GGA GAC ACA GGC AAG | chr3:88334036-88334485 |
| | | TCT CTA CAG CCA AGG GTC A | |
| | exon 2 | TTC CTC CCC ACA GTG TTC T | chr3:88333308-88333808 |
| | | AAT GGC TCC TCC GAC TTA G | |
| | exon 3 | CAG AGG CAC CTA CTA AGT CG | chr3:88332779- |

| | | | |
|---|---|---|---|
| | | AGT GGC AGG TTC AGG ATA C | 88333338 |
| | exon 4 | GGT TTC TCT TCA TTC TGC TG | chr3:88328481-88332829 |
| | | AAG GCA ATG GAT TCA ACA C | |
| | | GGT GTT TGG AGT AGT ATC GGC | |
| | | TTT TCC TTC TCC TTG GCG | |
| | | ACA CTT CCT TCT TGG GCA G | |
| | | GCT TGG AAG GCA GTT AGT TTC | |
| | | CAA GCC TCT CTC TGT CTT CC | |
| | | TGC TCC CAG TAA GTT GTC C | |
| DV045339 | | GAA ACA GAT GTG GAG GAC TTC | chr3:88689412+88690082 |
| | | GGC AAA GTC TCT CAG TGA AAC | |
| AI561824 | | TCT ATG GGC TTT AGG GAA TC | chr3:88692975+88693600 |
| | | TCT TGC CAA AAC CTA CTC C | |
| BB664182 | | TCC TCT TGC CAG ATA AGT CC | chr3:88696063+88696661 |
| | | AAT AAG AGC AGC CTG GAC C | |
| GH455320 | | AGT GCT GGG AAA GAG GGT AG | chr3:88704494+88705157 |
| | | TCT GGG GAA CTC TGG GAT AC | |
| AW764307 | | GCA CCT TTA GGC ACA CTT G | chr3:88712445+88713132 |
| | | TGT TGG GCA GGG TTT ATG | |
| gm15998 | exon 1 | AGA GAG TGT GTG GGC AAA G | chr3:89315287+89315605 |
| | | CAC AGC CCT TAC CAG TTT G | |
| | | AAT AGG CAA ACT GGT AAG GG | chr3:89315581+89316108 |
| | | AGT GTC TGA GGA TGA AGG TG | |
| | exon 2 | GGA AAG TCA GGG CCA GGA | chr3:89315920+89316688 |
| | | CAC TAC AAC AGC TCA GGG C | |
| | exon 3 | CCC AGG ATA GGC AGA GTA AC | chr3:89317762+89318313 |
| | | TTA CCC ATC TGC CTC GTT CC | |
| | | CCT GTC ATT ACC CAA CAT CG | chr3:89318239+89318953 |
| | | TGC CAG CAG GTT TTG AAC | |
| 4731419I09Rik | exon 1 | GGA ACT CAA CGT TCT TTC AC | chr3:89337450+89338099 |
| | | GGA AAG TTA CCC AGG GCT | |
| | exon 2 | CCC TAG AAA TGC AAA GAT | chr3:89345062+89347949 |
| | | GCC CCA TCA CTA TCT TAC | |
| | | AGG GCT GGA AGA ACT AAA AG | |
| | | TCT GTC CTT GGT CCA CTA TG | |
| | | GAA GAA GAC CTC ACA GGG | |
| | | ATG GAG ATT ACA CTA CAG TGT G | |
| | | GTG TTC CAG GCA CCA GTT CTA | |
| | | ATG TGG CTT CCC TGC TGT AG | |
| | | CTG ATC CCA GAG CAG GCA | |
| | | GCC TGA GTT CGT ACA GTT ATA G | |
| | | CAC ACA GAT GAG CTG TCG CT | |
| | | ACA GAC TGG TTT ACC TGA GGC | |
| 4632404H12Rik | exon 1 | CAG CGC TTC CCT GTC ACA T | chr3:89772220-89773323 |
| | | TTT CAG GAC CCT CCG TTG TC | |
| | | GGC ACT ATC ACA AAG GAC AAC | |
| | | ACC CTC AGC CCT TAT TTA TTA G | |

| | | | |
|---|---|---|---|
| | exon 2 | ACA TTG GAA ATC TGC CCG | chr3:89769042-89769650 |
| | | GCC AGG AGA AGC TGA GCT G | |
| | | CTC AGT ACT GGC AAC TCC | chr3:89768491-89769109 |
| | | GGA CTC ATT CTG TAG ACA GTC | |
| | | TCC CAG CAT TCA GGA AGC | chr3:89767996-89768559 |
| | | TGA TAC AGA TAC ACA CAC CGC | |
| 4933434E20Rik | exon 1 | GCT TCC TCT TAC TGG CGT A | chr3:90052744+90053225 |
| | | CTG AAG AAC AAA GAC AGA ACG | |
| | exon 2 | CCA GGG ATT TTG AAA GTG AAC | chr3:90053436+90053927 |
| | | TAA GCA AGC ACA ATG GGC | |
| | exon 3-4 | CCA AGC ATT TAC AAA CCA CAC C | chr3:90056074+90056640 |
| | | GCT CAT CTT TCT GTT CCA TAG G | |
| | exon 5-6 | ACA CGG GGT TTC TTG AT | chr3:90058397+90059343 |
| | | GTG TAG CTC TTC TGT ATT CTC C | |
| | | GTC TAC ACC AGA CCA GGC T | |
| | | AAA GCA CAA AGA GGA CCC | |
| 1700094D03Rik | exon 1-2 | TAG GGT CAT TGT TGG ACG CC | chr3:90062732+90063332 |
| | | TCA CGG AGA GCA GAT TTC GG | |
| | exon 3 | AGT CAG TTC AAA GCA GCG | chr3:90064355+90064821 |
| | | GCT AAA AGA GCA TCT CAC CTG | |
| | exon 4 | GAA GAG CCG ACA CCA AAC | chr3:90065470+90066017 |
| | | TGA CTT CAG ACA CTA GAA GAG G | |
| | | CCC ATA ATG AAA TCT GAT GC | chr3:90065976+90066664 |
| | | CTG AAC AGC AGC ACT ATC G | |
| | | CCC GGC AAT GGC TAA TAT TT | chr3:90066610+90067195 |
| | | TCA AAG CCT GCT TTC CCC | |
| | | GGC TTT GAA CAG GAG TTA CAA G | chr3:90067188+90067645 |
| | | TGC ATG GGT GTT TTA TCT GC | |
| | | GGA CCA GGC AAA CGC ATG | chr3:90067596+90068232 |
| | | GTT GAT TAA CTC AGG TCA TCG GG | |
| 2500003M10rik | exon 1 | AAA GTC AGT GGG GAT TCG | chr3:90508945-90509604 |
| | | TTT GAG GCC AAA CCA CTA C | |
| | exon 2 | ACT CAA AGG GAG CCA CCT AC | chr3:90505752-90507824 |
| | | TCT TGC AGC CAG GAT ATA CC | |
| | | TTT AAA AGA CTC CTC ACA CC | |
| | | CTT AGT AAC ACA TGA ATA CAG G | |
| | | GCA CAT GTG CTT CTC TGT TCA | |
| | | TCC ATC TGC TGG GCC AGT | |
| | | GAA CAA ACA GCC GAT GC | |
| | | AAG AAG TGA ACC TGG GAC C | |
| | exon 3 | TGG CTC TGC TGT TGT AGA C | chr3:90502894- |

| | | | |
|---|---|---|---|
| | | CCT CTC ACC ATT ACA CAT CAT | 90503521 |
| | exon 4 | GTA TGC CTC CTC TCC AAG GTC | chr3:90501797-90502271 |
| | | GTG ATG GTG GTG GCA ACA | |
| EL605591 | | TGA AAC CCC AAA GTC CAT AC | chr3:91054885+91055818 |
| | | GGT TTC CTT GGC ATA GAA TG | |
| | | TCT AAG ATA AAG TGG AAT CCC C | |
| | | TAT GGA CTC CCT TGC CTG AC | |
| 4930529C04rik | | GGA GGG CAT AGG AAA TGA AG | chr3:91071327+91075389 |
| | | ACA GAC AGT GGC GAT GAA C | |
| | | GAA AAT GGA AGT CTA TCG GC | |
| | | AAG GGG AAT CTC AAC TCC G | |
| | | CCA GTT GTT TGA TGC GAC | |
| | | CAC TTG ACC TTC AGC CAT AC | |
| | | ATT GGC ATC ACT GTC ACT AAC | |
| | | GAG AGC CTT TGG TCA CTT GCA C | |
| | | TAG TCA TCC TCA CTC TCG GC | |
| | | AGA GAG TGA GGG GAA AAG G | |
| | | AGG ACT CTT TGA CTT GTT TGC | |
| | | ACA ACT GGT GCC AAC ATC | |
| | | GGG TCA GGT ATT CAC GAG TC | |
| | | CAT CCT CTG GCT CTT GAT AG | |
| | | CTT CCT CCA AAT ACG CAA G | |
| | | TGG GCT TCT GAA AAC GCA CC | |
| AK078932 (9130204L05Rik) | exon 1-2 | TGT AGA CTT GGC AGG GTC AC | chr3:91090240-91090915 |
| | | CCA AAC TCA GGA ACA AGA CC | |
| | exon 3 | TCC ATA GAG AAA AAT GTG GG | chr3:91088071-91088728 |
| | | CAA CTG ATG CTC ATT CGC | |
| map0.2 | | CCT GGG CAA AGC AAG GA | chr3:91108492-91108826 |
| | | ACG TAT GTG ACA CAC CAC A | |
| map0.3 | | GTG TCA GCC TTT CTG GTA CC | chr3:91160352-91161039 |
| | | GTT TGG GCT GAG GCC CAG GAA | |
| map0.4 | | GCC TCA GCA ATC GCA GGT | chr3:91208775-91209359 |
| | | AAC AGC CTG GCA ACA AA | |
| map0.10 | | CAT AAG TGG T TA GCT TCT G | chr3:91510991-91511509 |
| | | AAT CCA TAA ACC CAC A | |
| map0.11 | | CAT GAA ACT CAA GAG GAA CGA AA | chr3:91561439-91561687 |
| | | GAA CTG GAA ATA ATC ACC CA | |
| map0.12 | | TGA CAT ACA TAG GAG GTT T | chr3:91608323-91608869 |
| | | TTG AGG GTC TAT GCC TTT G | |
| map1.5 | | CCA GGA AAA CTT ACA ACG GAA | chr3:91671237-91671835 |
| | | CTT CAA CAC TCC ACT CTC | |
| map1.11 | | GTT TGG GTA TGT TGT GGC TT | chr3:91733235-91733906 |
| | | CAA CCA TTT GCT TGG AAA ATT | |
| map1.12 | | TGT GTT GAT GTC TTC TAT GTT | chr3:91943487-91944191 |
| | | GGA AGC AAA ATA GGG CTC AG | |
| map1.1f | | GCA GCA CCC TCT GAA GAT GA | chr3:91996416+91996718 |
| | | GCT AAG CAC TTG TGG CTC AC | |

| | | | |
|---|---|---|---|
| map1.16 | | TAT GGC TGC CTC TAA GCT CCT G | chr3:92034176-92034817 |
| | | GAA TCT GGC AAT CAC AGC A | |
| map1.17 | | AAT TCT CCA ATA TTA ATG TAA A | chr3:92045514-92046222 |
| | | GAG CTG GGA AGT AAA GCA CA | |
| map1.18 | | GCT TCT GAA TTC TCA ATG CC | chr3:92054107-92054785 |
| | | GTG TCC CTC TTT TCC CCA CAT C | |
| CX208588 | | CCA TAA CCA GAG GGC TAT TC | chr3:92083395+92084085 |
| | | GGA TTC CTG CTT GGT AGA CC | |
| BB176842 | | CAG AAC ACT TTT CCT TCC AAT C | chr3:92084183+92085136 |
| | | ATC CTG GGT GTT CCT GTT C | |
| | | AAG TTG AGA AGC CTG CCC TC | |
| | | GGG GTC TAT TAG GGA CTT GG | |
| BB658268 | | AAG ACT GAA CTT TAT GCT GTG C | chr3:92097160+92098205 |
| | | GGT CAA CAG GGC AAG TTT AC | |
| | | AGG AAG ACC TTG AGG ATG AC | |
| | | GAG GAA GAA AGA CAA ACA AGG | |
| BB471038 | | CAT TGC TTC TAA CTC TGT GCC | chr3:92098830+92099879 |
| | | CAG GGC AAT AAT GGT GAA G | |
| | | CCA CCA GCA TTG AAA ATA CC | |
| | | CAG AAA AAA AAA CCC CAC C | |
| CJ061957 | | CAA AAA TGT TGT GGG CAA TC | chr3:92112866+92113510 |
| | | TGG GGA GAG AAG ATA GGA GTC | |
| map2.1 | | GGC ACC TTT ATT TCT CTC TGC | chr3:92121810+92122092 |
| | | CAA AAC ATC AGC AGA AAG ACC | |
| A030004J04Rik | | CCA CTC GCA ATC TTT CAC | chr3:92122894-92123443 |
| | | GAT ACC CAC AGG AGA AGG TC | |
| BY089468 | | CTT GCT CTG GGG TTC TAA G | chr3:91928538-91929000 |
| | | CCA TCA CTG GAA AGA GAG G | |
| EH097951 | | AAG CAC TAC ACA TCT GCC AC | chr3:92131423+92132720 |
| | | AAC AGG TTC CAT CTT GCT G | |
| | | ATT AGT TCT GGT GCC TGG G | |
| | | AAG CAA GAA ATG CCT CAA C | |
| lelp1 | exon 1 | CAT AGC TAG ATT CTC ATG TCC | chr3:92142466-92142877 |
| | | GGA GGC AGC TCT TTA AAT | |
| | exon 2 | GAG ATG GTC CAG AGA CTG AAC A | chr3:92134908-92135876 |
| | | TTG CCC ATG GTG CCT CAC TTT | |
| | | CCA AAA AGG ATG TGT CAA GC | |
| | | ATC TCT CAG TGC CAG GGA TT | |
| AK137696 | exon 1 | GTG AGC TGC AAC ATG AAT CA | chr3:92170151+92170546 |
| | | CCA ATC TGG CTT ATA TGA GA | |
| | exon 2 | TTC CCT TTC CTT GGC TGG TC | chr3:92170933+92171465 |
| | | AGT GAT TCT GTC GTG AGG GC | |
| | exon 3 | TGT GAC CTG AAA TGA GCC | chr3:92171998+92172492 |
| | | CTG GTA GTA GCA CCT TGA C | |
| | exon 4 | GTT TCT GGA AGC TTC ATT TGC C | chr3:92173482+92178122 |
| | | AGA TGG TCC GTC CTT TCG TC | |
| | | CTT GGT GAC TGT TAC CTA A | |

| | | | |
|---|---|---|---|
| | | CCC ACT GTC AAA CGA AGC | |
| | | GCA TGG CTA CCC AAC ATG A | |
| | | CCC TCT TAC TCT TAT ATA CCC C | |
| | | CCA CTG CAA CTG CCT GAA TC | |
| | | CAT CCA GAG ACT GCC ATA CC | |
| | | CAT CCA CAA TAG TGC CTG GG | |
| | | GGA ATG TTC CTC TTT CTC TAC | |
| | | CTA CCT CAC ACC AGT CAG AAT A | |
| | | GCT CTT GTG CAA CTC CAG TG | |
| | | GGT CCC TTG ATA GAA CTA | |
| | | CTT GAT GTA GTC CCA TTT GTC C | |
| | | CAC TGG ACA AAT GGG ACT AC | |
| | | TGT AAC AAA CAT GGC ATG CAG G | |
| map3.1f | | CAC TCT ACC ACC ACA ATC TGT C | chr3:92214798+ 92215068 |
| | | GAA GCA AAA GAT ACC AAG TGC | |
| AK033478 | | CCC CAT AAG TGT AAC CAC G | chr3:92256377+ 92258533 |
| | | GTC CAG CCT CAT CTA CAG AAC A | |
| | | CTA GTA CTA GCT GTT CTT GAA C | |
| | | AGC AAG AAA AGG AGG TTA GC | |
| | | AAC CTT ACA TAA CTT TAG CCC C | |
| | | CTG AAT CAT TGT ACC CCT CTT T | |
| | | TGT CAT TGT GAT AAG GGT CTC | |
| | | CAC CTG CTA TCT CTG CTC TG | |
| DT930973 | | TCT CAT CCT GCC CAA ACA C | chr3:92262832+ 92263669 |
| | | TAG GAC AGA GGG AGG TTA GG | |
| AK138172 | | ACA AAA ACA ACC CTC CCC CC | chr3:92282613+ 92283369 |
| | | GAG CCA TTG TTT TAG CAC C | |
| AK029022 | exon 1 | GCT GAG CAC TTT TTC ACG | chr3:92277470+ 92278889 |
| | | CAC CAA GGG GCA ACT GA | |
| | | GTT CAC CAG AGG CAA AAC AG | |
| | | TGA TAC CAA CCC TTG TCG G | |
| | exon 2 | TAA CTT CCT GAT TCC CGA C | chr3:92278857+ 92279831 |
| | | CAT CTC AAG TCC TTC TAT GAG G | |
| | | GCA TGG TAG GTA AGT TAC TGA | |
| | | GCT ACT GCT GTT TTC TGA CC | |
| | exon 3 | GCT CTG CAC AAC TCA TCT CTG | chr3:92279749+ 92280809 |
| | | GAT CCC TCT GTC TCC TAT AAT T | |
| | | CAA ACA TTA GGC AGA GCT CA | |
| | | GCA AGA CTT CCA GTG GAG GG | |
| | exon 4 | ACT CCC ACA GTT TTT GAG C | chr3:92280697+ 92281112 |
| | | AAT CAG GTG TAA CCG TGC C | |
| map3.1a | | TGC TGT CTT TTT CCC ACT G | chr3:92287510+ 92288072 |
| | | CAC AAA GAA TAA ACA GGC TGA G | |
| map3.1b | | CAG TGT TGG GGA AGA GAA G | chr3:92307494+ 92308026 |
| | | ATC CAG TCT CAT TGT CAA GG | |
| map3.1c | | AGA TTC TTA GGG AAA GTG GC | chr3:92317219+ 92317775 |
| | | TAC TTC CCA GTC CTG ATG AC | |

| | | | |
|---|---|---|---|
| map3.1d | | TGC TCT ACA TTG ACC CCT TC<br>CAT CCA CAT AAT AGC AGG ACC | chr3:92327483+92328016 |
| map3.1e | | GAT GGG GAT AGC ATT GAA TC<br>TTT CTG CCC TAA ATA CAG TCC | chr3:92337393+92337942 |
| map3.1g | | CAT TGT TTG TGA GTT GGT CTG<br>GAC TCC CTT CTT CCA ATG TC | chr3:92347434+92348065 |
| map3.1i | | TGG AAG AAA CCC ACA GTG<br>GTA GGC AGG AGT GGA TAG TAT G | chr3:92367468+92367920 |
| map4.1f | | ATG CCA GAA CTT GAG ATT GC<br>TAC ATC AAA ATG CTT AGG GG | chr3:92374334+92374609 |
| map4.1a | | TGT TAG CAG CAG TAA AAG CAG<br>AAC TCT CAG TGA TGG TGG C | chr3:92378450+92379024 |
| map4.1b | | GGC TCT TCT CCA TTT TTT GTC<br>TGG GAA AGA GGA CTG AGA GTA G | chr3:92388194+92388731 |
| map4.1c | | CCC TTT TAG TTG GTG AAT AGT G<br>CTG TTG GTA AGA CAT TGT GAT G | chr3:92398146+92398634 |
| map4.1d | | TTC TTT CAT TTG GCA GGC<br>GAG GAC AAG GTT CAG GAC AC | chr3:92408203+92408827 |
| map4.1e | | AGG TAG GTG GTT GAA CAG CG<br>TAG GAC TCT GGG CAC TTT GC | chr3:92418326+92418952 |
| map4.1g | | AAG AAG CCA TTT GTA GAG CC<br>ATT TAC CCT CAG GCA GGT C | chr3:92438343+92438838 |
| map4.1h | | GCT CAC AGG TAT GTA GGC AC<br>CTC CTG ATT CCT GGC TAT G | chr3:92448248+92448709 |
| map4.1j | | TCT CTC TTT CTT TCT CCT ACC C<br>AGT CTC AGG ACA AAT GAA GTT G | chr3:92468437+92468898 |
| map4.1k | | AAG ACA CAT TAG CAC ACA TCA G<br>CAG GTC CTC TAA AAG AAC ACC | chr3:92478331+92478814 |
| map4.1l | | GTT CTA TCC ATT TGT CGG C<br>GCT GGT GCC AAG TCT AAA G | chr3:92488339+92488988 |
| AK214076 | | GTG TGT AAA GAG GAA TCC CC<br>GGT TGT TCA GAC ACT CGT TG<br>CAA GAC TGA CCA AGA TGA GG<br>TGT GCT GGT TCC ATT ACC<br>GGT AAT GGA ACC AGC ACA G<br>ACC ACT ACG GAA TCA CAC AG | chr3:92428114-92429476 |
| sprr3 | | AAG ACT TCA GCA GCA GGT G<br>CAC ACA GAT GTT CCA CAA CC | chr3:92456763+92457409 |
| map5.1 | | GGG GCA GTG GAT TAT TTT CT<br>GCT GGA GTT TCT CAA GTT TTG | chr3:92504884+92505256 |
| map5.1a | | TGA CTG TTG TCT GTG AAT CTT G<br>TCT CAC CAA CAG ATA AGT CAC C | chr3:92514263+92514825 |
| map5.2 | | GCA GTA GAT GAA AA GCA GAA TC<br>GAG AGG GTC AAG TA TGTC AGC | chr3:92509258+92509814 |
| map5.3 | | TGG TAT GGT GAC AAA CAG G<br>TTT TCA GCA GTT CGG TAT TC | chr3:92519331+92519939 |

| | | | |
|---|---|---|---|
| map5.1b | | AGG TGT GTG GCT TCA TTT C<br>CAC AAA GAA TAA ACA GGC TGA G | ch3:92328000-92329000 |
| map5.4 | | TAT CAG CAG ACC TGG GAG AC<br>GCC ACC AGG TAG TTT TAT GC | chr3:92529377+92529933 |
| map5.5 | | CTG TGG CTC AGA CTT TTG AC<br>CTC AGC ACA TAG CAG GTA TTC | chr3:92539187+92539710 |
| map5.1c | | GCC TTA GAG CAA ATG TTT TCA C<br>ATT AGA GAC CAG GAG CCA GG | chr3:92544237+92544699 |
| map5.6 | | CCA CAC TTC AGT CTT CAT TCT C<br>GAT GTG GAG AAA GAG GAA CAC | chr3:92351414+92351923 |
| map5.1d | | AGA CTC TGC TGG CAA GGT AG<br>TCC TGC TGA CTG CTT GAA G | chr3:92554250+92554762 |
| map5.7 | | TTA GGG AGT TTT CGG AGG<br>CCC TGT AGG AAC TAT GAC CTC | chr3:92559416+92559980 |
| map5.1e | | CTC ACA GCC ATC CAT TAG AC<br>CTT CTG TAG TGT GGC TTA TGG | chr3:92564480+92565076 |
| map5.8 | | CAC CTC CAG CAG AAA GAT G<br>GAA TGA ACA CAG ACC CAG C | chr3:92569253+92569764 |
| map5.1g | | TGG TTC ACT TCA GGC AAA G<br>TGA ATC TTT GAA TCC AAG GC | chr3:92574388+92574913 |
| map5.9 | | GCT TAT TAG CAG AGA GGG TTG<br>AGA GTG AAG TTG GAG CAC AG | chr3:92579183+92579752 |
| map5.1h | | AAG AGC AGC AGG TAG GTT G<br>TTA CCC TGG AGA TTT AGA TGC | chr3:92584229+92584721 |
| map5.10 | | TCT GTA TGC TCA TTG TCC AAG<br>TTG AAA CCT CAC CTG AAG TG | chr3:92599528+92600064 |
| map5.11 | | CAA CAC TGA AGA AGG CTG G<br>CAA AAG TCA TTC AGA TAG CCA G | chr3:92610322+92610938 |
| map5.1i | | TAA GGA GGG GCA TAG TGT C<br>AAA TGC CCA CAT CAT AAC AG | chr3:92614140+92614720 |
| map5.12 | | CCT GGG ATG AAA TCT ACT TGG<br>GTG TTT TGT TCC CAT TGC TAA G | chr3:92626325+92626880 |
| AK019072 | | TGA GTG CTA ATG CTG GCT C<br>TTG AAC TGA AAG GGT AGG C | chr3:92620109-92620641 |
| map6.1 | | GGC ATC AAA TGA AGT TAG CA<br>TCC TAA TGT CAA GTT CAC AGG | chr3:92631383+92631730 |
| map6.2 | | GGA AGG GAG CAT TTA TTT ACT G<br>TGT TCT TCT TCT CCA CCC C | chr3:92635279+92635906 |
| CR514969 | | CCT CTA TTT CTA TCG GCA CC<br>CCT TAT TCC TAC AGA GTG AGA A | chr3:92638914+92639499 |
| map6.3 | | CGA TGG AAA AAC CTC AGC<br>TCT TCT AAG CCC TCT AAG CG | chr3:92640211+92640613 |
| map6.4 | | GCA AAA GGA TGT GTA GCA AAC<br>GGG TTA GGA ATG TGG TTC AG | chr3:92645330+92645883 |
| map6.5 | | CGT CAT TTA TTG TGT GGG AG<br>CAG GTA AGG AAA TAC AAA ACC C | chr3:92650501+92650898 |
| map6.5a | | CTT GCT CTT CCT CAG GTT G | chr3:92651084+ |

| | | | |
|---|---|---|---|
| | | TGA TAC TGG CTC AGG AAC AC | 92651515 |
| map6.6 | | CCT CGT GTT CTA AGG TTT CC | chr3:92655081+ 92655658 |
| | | GCA AAC ACG AGT AAA CTT GTG | |
| map6.7 | | CTC AGC ATT CTT TAG TTG CC | chr3:92660587+ 92660981 |
| | | CCA CAG ATA AGT GTC ATC CC | |
| map6.8 | | GTT GTA AAA GAA CTG CTG AAG G | chr3:92665278+ 92665808 |
| | | CTT CCC ATT CTG TAG GCT G | |
| map6.9 | | GGA TGT GTA TGA ATG GAG ATG | chr3:92670302+ 92670743 |
| | | CTC TTT GAC TTT CAC AAG CC | |
| map6.10 | | CTG ATA AGC CAT TGT TCC AC | chr3:92675212+ 92675834 |
| | | CAT CCA CCA GGA ACT AAC ATA G | |
| map6.11 | | AAA ATG CCC CCC TAA ATG | chr3:92680360+ 92680878 |
| | | GAA GAA AGG GAT TAG GTC ATT G | |
| map6.12 | | GCA AAG AGG CAT TGT CAT C | chr3:92685533+ 92686002 |
| | | CCT AAG TGT CCT CCC AAA TG | |
| lce1d | | AGG CTA CAG CAG GAA GAC AC | chr3:92685956+ 92686594 |
| | | CAG ACA TTT CAT TGA GGG AAT C | |
| | | GTG ATT CCC TCA ATG AAA TG | chr3:92686571+ 92687111 |
| | | GAT GAG AGA TGG AGG AGA GAC | |
| AK202131 | | GCT GGT GAT GAC AAG AAA GG | chr3:92689825+ 92690352 |
| | | CTA AAG TTA ACT GGC TCT ACT C | |
| map6.13 | | CTG GAC CAC TTC ATT TTC TAT C | chr3:92691529+ 92691934 |
| | | GTT CTT TGG AAT AAC CTG ACC | |
| map6.14 | | TCT TGT GTT TCC TTG GAT GTA G | chr3:92696426+ 92697065 |
| | | CAG AGA AAA GAG AAA CAG AGG C | |
| map6.15a | | GAT GTG TTC TTG GAT TCA GTG | chr3:92702374+ 92702768 |
| | | AAA ACC TGG ATG AAA TGG AC | |
| map6.16a | | CCA TTT CAA AGT GTT CGC | chr3:92707188+ 92707595 |
| | | AAG ACT TCA GAC AAA TGG GAC | |
| map6.17 | | GAA TAG GAA GGA TAG CAG GC | chr3:92710134+ 92710651 |
| | | TTG GGT TTC TTT GGG AAC | |
| map6.18 | | TTG TGG CTA TTG TGA CTG G | chr3:92715566+ 92716077 |
| | | ACC ACC AAA CCC AAA GAC | |
| map6.19 | | TAG TCA CCA CCT TTC CTC C | chr3:92720485+ 92720975 |
| | | GGA TGC TTT ATT TCT GTG AGA C | |
| map6.20 | | TGT TTG GGG ATT TCT ATG G | chr3:92725231+ 92725851 |
| | | TCC TTT CCT CCA GTT TAT GTC | |
| map6.21 | | CAA GTC AGG TGA AGC AAG C | chr3:92730543+ 92731077 |
| | | GGG AGA GAA GAA GCA GAC ATA C | |
| map6.22 | | GCA AGG AGG AAA ACA AAT G | chr3:92735431+ 92736027 |
| | | GGT CCA GGT TTG TTG ACT C | |
| map6.23a | | GGA CTT GCT CAG GAC TCT ATG | chr3:92741101+ 92741593 |
| | | TGT AAT AAC CTG GGC AAG AG | |
| map6.24 | | ACT GTG GGC ATC ACT AAA AC | chr3:92745481+ 92746009 |
| | | GGT TGC TTC TGT ATC CTG G | |
| map6.25 | | ACA AGA AAC AGG ATG GGC | chr3:92750319+ 92750944 |
| | | AGT GTC AGA CCC CAA AGT G | |

| | | | |
|---|---|---|---|
| map6.26 | | GGA GAG GAA AAG TAT GCC TG<br>GAT AAG CAT TTG GCA CTG G | chr3:92755301+<br>92755898 |
| map6.28 | | AGT GAC TGG AGA CAG GAT GC<br>CCA CCT TTC AAC TGA TGT CTT C | chr3:92752401+<br>92753043 |
| map6.29 | | GCT AAA AGC CAA AGG TGG<br>TGT GTT GTG TTT GGT GAG G | chr3:92757246+<br>92757689 |
| map6.27 | | CCC TGA GAT TCT GTC CTA CAA C<br>CCT ATG ACT ATT TTG TCC CCC | chr3:92761467+<br>92762073 |
| map7.1 | | GGT GGT GTC CTG TTG AGT GA<br>GAG GCA ATG GTG TCT GTG AC | chr3:92764574+<br>92764879 |
| map7.2 | | GCC AGA AGT TCA GTG AGT TG<br>AAA GAG AAT GTT GGT GTC TGA G | chr3:92769112+<br>92769697 |
| map7.3 | | CAG AGA CTG AGC CAC AAA TC<br>TAT TGG ATG GGT CAC ACG | chr3:92774141+<br>92774747 |
| map7.4 | | AGG TGC CAT TCT GAA AGT C<br>CCA TCT CCA TCC ACT GTT TAG | chr3:92779369+<br>92779970 |
| BY708731 | | AGG GTA AGG GGT CAG TAT GG<br>AAG TCT TCC TCA CCC AGG TC<br>ATC TCT CCA TCG CCA TCG TC<br>CAA CTC CCT GGA CAT TGA G | chr3:92788605-<br>92789555 |
| map7.5 | | AAC TAC CCC AAC TAA TGA GGT C<br>TGG TAT GCT CTT CTC TGG G | chr3:92793078+<br>92793591 |
| BY708617 | | TGA ATC AGA ACG GAC AAG AG<br>TGA AAC ACA GAA GGG GTA GC | chr3:92806321-<br>92806971 |
| map7.19 | | TTC CTA CCT CAA CTT GCC C<br>GCC AAC TAC TCC TGG ACA TAA G | chr3:92809105+<br>92809692 |
| map7.6a | | CTT GAC TCT CTT TCA GAC TTG G<br>GGC ACT GTT GTC AGG AAT AG | chr3:92812216+<br>92812643 |
| map7.7 | | GCC CAG AGG AAT AGC ATA ATC<br>GGT TCT CAA CTT GTC TAC AGG G | chr3:92816238+<br>92816859 |
| map7.20 | | GGC ATT GTC TCC CTT TTT AC<br>CAT CCC TGT CTC TCC AGA TAC | chr3:92819267+<br>92819766 |
| map7.21 | | GAT TGC TGA GTT GGC TTT G<br>AGT CAG AGT TTG AGC AGT GG | chr3:92922996+<br>92923418 |
| kprp | | GTG AGA ATC TGA GGC AAG G<br>CTT CTC CAG GGA CAA CAA G | chr3:92826953-<br>92827583 |
| map7.23 | | CTC TAT TGA AGG AAG TCA GGG<br>GGG TGT TCT AAA GTG GCT TAT C | chr3:92836391+<br>92837040 |
| map7.9a | | CTT CCT GGG CAT TTT AGC<br>AAC CTT TGA GTT GGT GCT TAG | chr3:92838221+<br>92838869 |
| map7.24 | | AAC CTA TGC TAT GGC TAC CC<br>CGT TCT TCG TTC TTC TTG AG | chr3:92942557+<br>92943169 |
| BG068233 | | CAC AGT CAG GCA AAA AGG<br>TTT TCC CTT CTT TTC CCC | chr3:92843226+<br>92843879 |
| map7.25 | | GTT CTT GGC AAA AAA ACC TC<br>GCT GTT ATG TGG AGG AAC G | chr3:92846516+<br>92847165 |
| map7.10a | | AGG AGG AAG TTT GTT GTC ATC | chr3:92847315+ |

| | | | |
|---|---|---|---|
| | | CAT CAT TAC CTG CCC AGT AG | 92847841 |
| lce1l | | GAA AAA CAC CAT ACC CCA TC | chr3:92849620+ 92850065 |
| | | TCC CTC CCT TCT TGT TTT C | |
| | | AAT AGC AGG CAG GAA CAG G | chr3:92850068+ 92850659 |
| | | AGT TTG TCT TGG AAT CAG TGT C | |
| | | AAA GCC ATC AGC AGA AGT ATC | chr3:92850587+ 92851104 |
| | | GCT GAA GTT ATC TTT GCG TAT G | |
| | | CGA GAT GAG ACT GAA AAA CAA C | chr3:92851111+ 92851673 |
| | | CAT TGT AAA CCA TTG CTT GC | |
| map7.11 | | GAG GTT ATT ACA GGC TTT TGC | chr3:92852163+ 92852762 |
| | | GAT GCT TGA ACT TCC TTG G | |
| map7.26 | | AGG CTG CTA TGA ACC AAA G | chr3:92854300+ 92854873 |
| | | TCT TTT ACT GAA CGG CTG TC | |
| map7.12a | | ATG CCT CTA TGG GTT GTA TG | chr3:92856326+ 92856896 |
| | | GGT GAA TAC TCT TGA CTG CTG | |
| map7.27 | | TGA AGG AGG GTT GCT ACT C | chr3:92858303+ 92858967 |
| | | CAG AAC AAT GGC AGT ATC AAC | |
| map7.13a | | TCT CAC GCA GCA GAT GTT C | chr3:92861286+ 92861775 |
| | | AGA GGG AGA GAG GAA ACT GG | |
| 2310050C09 Rik (AK009900) | | AAA TCC CCA GCA GGA ATG | chr3:92869995- 92870597 |
| | | CCT GGT TCT CTT GAA GTC CC | |
| map7.14 | | CAT AAC TCT GTG ATG GGG C | chr3:92871177+ 92871711 |
| | | CCT TTA CTC CTG ACA GAC AGT G | |
| map7.15 | | CTT GCT GTG GAC TTC ACT G | chr3:92876098+ 92876700 |
| | | GGA CCA GAA AAG AAA ATC CTC | |
| map7.16 | | CCT TTG TGT GTA GTT GAA GCC | chr3:92881386+ 92882074 |
| | | GAA AAC AAG GAC AAA TAC ACC C | |
| map7.17a | | CCC CTG ATG TGA AAA GAA C | chr3:92885263+ 92885755 |
| | | GAA CAG AGT ATG GAA GAA GCA G | |
| map7.18a | | AAG AGG TTT CAA GTC AGA TGC | chr3:92890347+ 92890953 |
| | | CTC TGT TGG ATG TAG GTT GG | |
| map8.1 | | AAG CCT CCA TTT TCT CTA TT | chr3:92893942+ 92894275 |
| | | AGG TTC TTT GGT TTC TCT TAG C | |
| map8.2 | | GCA TAG CCT CAC AAG AGA CAG | chr3:92745902- 92746505 |
| | | TCC AGC AAT ACC TCT CCT G | |
| BY364870 | | AGT TTT CAT CTG TGC CCC | chr3:92901946+ 92902595 |
| | | TGG GTT ATC GGT GAG TGA G | |
| map8.11 | | CTA CAC TCC ACC TCA GTT TAC C | chr3:92906205+ 92906876 |
| | | TTC TCA GGC TCA CAG TAG TTG | |
| map8.12 | | CAT CTC ACT ACC TTG CTC ATT G | chr3:92911078+ 92911606 |
| | | TGA CCT GGT TCA CTT CTT CAG | |
| map8.13 | | GGG TCT GGA AGT CAA ATA GC | chr3:92916081+ 92916685 |
| | | AAA TCT GCC TGG AAG TGG | |
| map8.23 | | CTG GTG AGT CAG TGT TCT GTC | chr3:92918370+ 92918948 |
| | | ACA CTG ATG GAT ACT CCT GAA G | |
| map8.14 | | CAT CAC TGT CTA CAT CTG GGG | chr3:92921099+ |

| | | | |
|---|---|---|---|
| | | AGT GGG AAA AAA GGC AGC | 92921642 |
| AV087208 | | TTC CCA CAG AAG TCC TTT G | chr3:92925074+ 92925713 |
| | | CCA AAG AGC ACA ACA CAG TG | |
| CJ238734 (lce3a) | exon 1 | GCT TGT GAA TGA ATG GGG | chr3:92925888- 92926495 |
| | | CCT GAC CTA TCT GCT TCC TAA G | |
| | exon 2 | GAA GGA GTG TTT GGC TTA CAG | chr3:92925421- 92926073 |
| | | TGA AGA AAG AAC AAG CCA AC | |
| | | CTT CGT GCT GTG CTA CAA G | chr3:92925109- 92925679 |
| | | AGG CGT TAT TTC TTC ATT CC | |
| map8.15 | | CCC AGG AAA ACC AAA AGG | chr3:92930080+ 92930594 |
| | | AAG GCA AAG GTC AGC AGT C | |
| AK009001 (lce3b) (EH107862) | | GAT TAC CAG ACA GAC TCA GCC | chr3:92933392+ 92934019 |
| | | TGT GAA CAT CAG ACA ATC AGA G | |
| map8.16a | | GAT AAC ACA CTC CCT ACA GGC | chr3:92938646+ 92939248 |
| | | TCC TGG TAG TTA GCA TTA GTG C | |
| map8.22 | | ACC AAA CCT ACA TTG TCA GC | chr3:92941980+ 92942583 |
| | | CTT TGC TGG GTA GCC ATA G | |
| AK009306 (lce3c) | | AGT TTG GGT GTG TCT GCT C | chr3:92944200+ 92944764 |
| | | AGG TAA CCA CAA AGG CTC C | |
| | | ATG CCT TGC CAG ACA TTG | chr3:92944949+ 92945647 |
| | | CAT CAG ACA GGC AGA GAA GAC | |
| AV081052 | | ACA GTC TGG AGG AAT CAG C | chr3:92945555+ 92946178 |
| | | CAT TTT CCT GAA TCC AAG C | |
| U6 | | CTG ACT GTC TGT CCC GTT AC | chr3:92949020+ 92949653 |
| | | TCA GAT TTC TTT ACG GAT GG | |
| map8.3 | | GCT CCT TGT GTG ACT TTC AG | chr3:92950406+ 92951071 |
| | | GTT TTA GCA AGG CGA AGG | |
| map8.1a | | CTT AGG CAA ATG GGT GAA C | chr3:92955474+ 92956027 |
| | | CTT GGG TGG TGG TTT AGA C | |
| EH095282 | exon 1 | GGC AGA CCT GAG TGA AAA GTC | chr3:92957193+ 92957595 |
| | | AGG CTC TTG TCC ATT TCC C | |
| | exon 2 | TTA GGA AGC AGA TAG GTC AGG | chr3:92957961+ 92958649 |
| | | TGA GTT AGA GGG GAA GAA CAC | |
| map8.21 | | CAC AGA GGA AAT CTT ATC CAA C | chr3:92959255+ 92959864 |
| | | AGG GCT ATC CTG GTT TCT AC | |
| map8.4 | | TCA CCA CAC AGT AGC CTT G | chr3:92961386+ 92961901 |
| | | ATA ACT TGC TGT GTA ACG GG | |
| map8.20 | | GAA AAA GAA GAA GGT TCC TCA C | chr3:92963230+ 92963740 |
| | | TGG CTG CTC TTT CAG AAT AC | |
| map8.1b | | TGC CCT TCT TCT AAT GAT GC | chr3:92965475+ 92966068 |
| | | GCA CAA AAA AAG GAA CGA G | |
| map8.19 | | GTC TTT TCC ATT CCC AGG | chr3:92968108+ 92968706 |
| | | GAC CTT CAT CAA CAG TCC C | |
| map8.1c | | AGG CAA ATA GAT GGA CCT G | chr3:92975472+ 92976018 |
| | | CTC CAG GGT ATT GGT TAG TTC | |
| map8.6a | | CAC TCC AAA AGA ATG TGT GTC | chr3:92980111+ |

| | | | |
|---|---|---|---|
| | | AAA GAA CAC AGA GGT GGT TG | 92980573 |
| map8.18 | | GAA AGG CTT GGG ATT TAC TG | chr3:92983186+ 92983867 |
| | | AGC AAT GTC TGG GAG AAA TAG | |
| map8.1d | | AAA GCC AAA GGA CAC TGA C | chr3:92985482+ 92985972 |
| | | AGT CAA CAT AGC AAA GGG ATA C | |
| map8.7a | | GCT GCC TGT TGA TTG TTG | chr3:92991193+ 92991762 |
| | | TGT GGG AGG CAG AAC TAA C | |
| BC060281 | exon 1 | TGA GAA TCA GGA GCA AAT CAC | chr3:92991840+ 92992377 |
| | | CAC TTT CTG CTG CCT TCA C | |
| | exon 2 | GCC TTT GTA AAC CTC GTT G | chr3:92992649+ 92993368 |
| | | CCA GGC AGT TAT CAA AAG C | |
| EH105059 | | AAC CTC AAA TGT TCA GTC CTC | chr3:92993219+ 92993734 |
| | | CCT TTA CTT CCT TAC TCT TGG C | |
| map8.8 | | GAC CTG GCA GTG AAT AAG G | chr3:92996415+ 92996894 |
| | | TCT TGA AAT ACA GCA CCT CC | |
| CB841495 | | AAA GGC TCC TGA GAA GTC C | chr3:92998924+ 92999462 |
| | | CTA AGC GTA GGT GAA AGC C | |
| | | CCA GGT TGT CTC AGG AAA C | chr3:92999347+ 92999940 |
| | | TCT CTT TTT TTC TCC CGC | |
| map8.9 | | CAG GTA ACA TCC CTC AAG TG | chr3:93002148+ 93002689 |
| | | GTA ACA TCT TTT GTC ACC TTC C | |
| map8.1e | | TCC AGA GAG AGG AAG TTG AAC | chr3:93005494+ 93005996 |
| | | CTG TAG CAA TCT GGA ACT GG | |
| map8.10 | | GAG GTC AAA GCA AGA TGT CTA C | chr3:93010342+ 93010849 |
| | | CTC CTT GAT TTG GTA GAA CTT G | |
| map9.1 | | TCC CAG AAG CAT TCA TCA CT | chr3:93015786+ 93016126 |
| | | CTA TGA AAG ACA AGA ACC AAG G | |
| crct1 | exon 1 | GAC CTG GAC TTG TTG GAA TC | chr3:93015484- 93015737 |
| | | ATA GGC AAC ACA AAA GGC TG | |
| | exon 2-3 | GCC AGA CAG ACTTT TCT CTA AG | chr3:93014444- 93014947 |
| | | ACT ACT ATG GAG GAC TGG GTG | |
| | exon 3 | GGA AGC CAA AGG TCT CAA C | chr3:93014104- 93014585 |
| | | GTG TTT GTT CTC CTT CAG GC | |
| CB590687 | | GAA AGG GAC TAT GAC TTG GTG | chr3:93017735+ 93018560 |
| | | AAC CTC TTC CTG CTG TAG CC | |
| map9.1h | | AAA GGA GTC CAT TCA ACC C | chr3:93027320+ 93027864 |
| | | GGC AGG CTG ATA CTT CTC TC | |
| map9.1i | | GCA AAG ATT CTA TGC CCC | chr3:93037487+ 93037978 |
| | | CAC AGC CAA ACT TTC TAA GC | |
| map9_3 | | CTC CAT TTC TGG TAG G | chr3:93041572+ 93042589 |
| | | ATC ACC ATC CCT GAC CTC | |
| map9.1j | | TGA AAC CTG TGT CGG AAG | chr3:93047457+ 93047942 |
| | | CAA GAT TCC CAA GTG TAT GG | |
| AV159627 | | CAG GTT AGT TGC TTC TGT TGG | chr3:93063361+ 93063894 |
| | | TGT GAG AGG GTA TCT GGT TTG | |
| AA261290 | | GAA AAG GGA TGT AGA AAT CAC G | chr3:93064325+ 93064848 |
| | | GAC TCC TAC AAA GAC AGC ACC | |

| | | | |
|---|---|---|---|
| gm4858 (EG229571) | exon 1 | CTC ATC TGG ACT ACC TCT TG<br>GAA CTT TGT ATG AGC TTG G | chr3:93068567+ 93069122 |
| | exon 2 | AAT GCC AAG GGT GAA AAA TG<br>GCT TTC TTT CCC CTT TAC TTC | chr3:93796683+ 93797909 |
| map9.1L | | GAT GAA CCC AGG AAA ACA G<br>AAA CAC CTG AGT AGA CGG G | chr3:93086127+ 93086626 |
| BC112282 | | CTG TAG ATA AGG GAT TCT GGG<br>TTT GTT TTT GCC ACG AGG | chr3:93100193- 93100744 |
| AV208954 | | TAT GGT TCC CTC AGA AGC C<br>CAA GAT GGA CAC GAA CCA C | chr3:93099783+ 93100385 |
| AV132146 | | CTT AGC AGG AAT TAA A<br>CAG ACA AGC TGG ACA TG | chr3:93099990+ 93100593 |
| AV208954 | | TAT GGT TCC CTC AGA AGC C<br>CAA GAT GGA CAC GAA CCA C | chr3:93099783+ 93100385 |
| BY702609 | | CAG GGT TTC TCT GTG TAG CC<br>TGT ATT GTG TCC ACC CTT CTC | chr3:92904976- 92905360 |
| map9.1a | | CAA TGC TGG GAA AGA AAA AG<br>GCA AAG GAG AAC CAC TGT C | chr3:93103091+ 93103656 |
| CK377705 | | TGG TCT CTC AGG AAC TAA AAG G<br>AGT TGC CAG AAA CCG ATG | chr3:93113090+ 93113583 |
| map9.1b | | ACT CTT GCT CCT CAG AAG CC<br>CAG AAG AGT CCG CCT CAT AG | chr3:93114950+ 93115532 |
| map9.1c | | AGT GGT TTT GTT TTG GGC<br>TTA CTG TAT CAT CCT TTC GCT C | chr3:92929902- 92930410 |
| map9.1d | | GAT GTT CCT GCT GAA GAC C<br>CCT TTT GTC CTC CCT TTG | chr3:93136063+ 93136529 |
| CRNN | intron 1 | ATG GAG GAA GAG GCT ATG C<br>ACA AAG GAA GAC GGA AGC | chr3:93145202+ 93145707 |
| CJ189654 | | GAC TTT TCC CGA TTC CAG<br>TCG CAT ACA TTC TTT CAA CC | chr3:93149470+ 93149936 |
| map9.1e | | ACA ATG AAA ATG GGG CAG<br>GCT TCT CCT CTT CCT ATT CAA G | chr3:93156135+ 93156743 |
| BY765199 | | TCT GGC TGG TAA AAA GGT C<br>TTC ACA ACA TCC TGC CTC | chr3:93173490+ 93174312 |
| BY750919 | | TGG TAG AAG TTG GGA TGA TTA G<br>GCC CTG AGT TTG ATT CAT AGC | chr3:93178228+ 93178845 |
| | | TGT CAG AAA TGT AGT GGT TCA G<br>CCT TCC AGA GAT GCC TAC TTC | chr3:93178715+ 93179199 |
| map9.1g | | TTT GTT CTC ATT TAC TGC CG<br>ATA GGT CAC TCT CTC CAA ACT G | chr3:93187985+ 93188532 |
| map10.1 | | CCACCTTCTTTGAATGCTAT<br>AAA GGA ACA AGT AGG ATA AGG C | chr3:93196979+ 93197318 |
| FLG2 | exon 1 | AAA TGA ATC AGC CCA TCC<br>CAA AGG TTG AGA ATC ACT GG | chr3:93197201+ 93197567 |
| | intron 1 | TCT GGA AGG ACA ACT ACA GG<br>ACG ATG TCA CTC CAA CGA G | chr3:93277680+ 93278101 |

| | | | |
|---|---|---|---|
| | exon 2 | AGG GGA GTC ACT ACA ACA TAA G | chr3:93200023+ 93200475 |
| | | AGA TTT ATG ATG ACC TAG CCC | |
| | intron 2 | TGA GAA AAC CCC AGT AAT AGT C | chr3:93200429+ 93200978 |
| | | ACG CCT ATG CTT CTT TGA C | |
| map11.2 | | CAG GAA GAG TAG AAC CAG AGG | chr3:93220670+ 93221181 |
| | | CTA AAA GTA TGT CCC TTG CTT C | |
| map11.2a | | GCT CAA AGC ACA CAT TGC | chr3:93228009+ 93228536 |
| | | TTT TTG GAG GCA CTC AGA G | |
| SF3A | | AGG CAT TAT CAT GAG GAG AAG G | chr3:93234942+ 93236468 |
| | | GAG CAG CCC TTG TCT TTC CAG | |
| | | GGG GCT ACA ACA GAA GAT CAG A | sequencing primers |
| | | GTG TGC CAA TGT CAA TGG | |
| | | TGA AAG AGA CAA GCA GCG | |
| | | CAG TGA GAG CAA AAG TGA GG | |
| | | TCT TCA CTT GGA TTC TCT CG | |
| | | GGA GAA TCT GGA GAG GTC AAG | |
| | | TCA TCC TCA CTT TTG CTC TC | |
| | | AAT CTG CCA GCG TGT CTT G | |
| map11.2b | | AAG GGC ATA AAT GGG GAG | chr3:93242063+ 93242675 |
| | | GTT GGG AGT GAG TTT TCT TTT C | |
| map11.2d | | AAG GAG GAA ATC TTG GGG | chr3:93262312+ 93262904 |
| | | CAG ACC TCT ATC CAA TGC TTG | |
| map12.1 | | CCC CAT CTT CCT CAA TAC CC | chr3:93272858+ 93273302 |
| | | TGG TGA GTT TCA GTA AGG GTC | |
| BY722912 | | TGA CAT TGG AGA AAT AGA GGT G | chr3:93273389+ 93273912 |
| | | GGC ATA CTT ATT GGT TGT TTG C | |
| | | TTC TTT GTC CAG TTC TGC TG | chr3:93273791+ 93274389 |
| | | ATC ACA GAG TTG ACT GGG C | |
| FLG | intron 1 | CAC TGA GAT TCG GGC TAT G | chr3:93278425+ 93279043 |
| | | TTC CTT AGA CTC TTC CTT TGC | |
| BY787539 | | AGT CCG ATT CCC AGC ATC AG | chr3:93282124+ 93282872 |
| | | TCC TCG CTG TGT TCT TGC TC | |
| AK076261 | | AGT CCG ATT CCC AGC ATC AG | chr3:93282124+ 93282872 |
| | | TCC TCG CTG TGT TCT TGC TC | |
| map13.1 | | CAG CCT CAG TCA CAC CTT TA | chr3:93317570+ 93317903 |
| | | TTC TCT GCT CAT TTT TGT GG | |
| map14.2 | | GTT TCT GGC TCA GGT CAC TC | chr3:93323174+ 93323599 |
| | | ATT TTG TCC TGA AGT GGA TG | |
| CN689430 | | CAA CAG GAT TGG AGC CTA C | chr3:93352513+ 93353252 |
| | | AGT AAT CTG CCT GCC CTT G | |
| RPTN | exon 1 | CCA ACC AGA GTA GAC AGT AAG G | chr3:93394632+ 93395345 |
| | | TTA TCA CGA GGA AGA GAA CAA C | |
| | exon 2 | TTT GGA GAC ATC CTT CGG | chr3:93395099+ 93395624 |
| | | TTA TGA TGG CAG GCT TGG | |
| | | TGG AAC ACT TGG ACC GAG AC | chr3:93395534+ 93396224 |
| | | CTC TCT GTC TGA CTG TTG CC | |
| TCHH | exon 1 | TGA AAT GAA AAC AGG GGG | chr3:93442034+ |

| | | | |
|---|---|---|---|
| | | TCA TCA GTG GTA ATG CCC | 93442722 |
| | intron 1 | GGT CTT GAT GGC ATT GTT AG | chr3:93442662+ 93443265 |
| | | GCT GGA GAG ATG GTT TAG CC | |
| | exon 2 | TAC AGA TGG GGT TGT GAG CCA C | chr3:93443165+ 93443748 |
| | | AGT CTC TGC TCT TCA TCG C | |
| map15.1 | | GCC CAG TTT TCT CCC ATA GTC | chr3:93463077+ 93463519 |
| | | AAT GTT TGT GTG GGT AAT GC | |
| BY367048 | | TCT CAG TGG CAA TGG GGT AG | chr3:93459816+ 93460670 |
| | | CAC AGT GCC AAA GGA ATG | |
| S100A11 | exon 1 | ACC TAC TTG ATG AGC ACG AC | chr3:93520213+ 93520739 |
| | | GTA ATT TTA GGG CTT CCT CG | |
| | exon 2 | TGG AAT AAA AGG GCA TAC G | chr3:93523933+ 93524682 |
| | | TCT TGC TCA TTC TGT CTT TAC C | |
| | exon 3 | GGA TTT GTG TCC AGG TCT G | chr3:93525849+ 93526418 |
| | | GAT TCT GGT TCC ATC GTT C | |
| DV066316 | exon 1 | ACC TAC TTG ATG AGC ACG AC | chr3:93520213+ 93520715 |
| | | CTC CAC TCT TCT CTA ATT GCC | |
| | exon 3 | GGA TTT GTG TCC AGG TCT G | chr3:93525849+ 93526418 |
| | | GAT TCT GGT TCC ATC GTT C | |
| BX518736 | | GAT GTC TGG GCA CAA ATG | chr3:93528637+ 93529287 |
| | | GGA AGA AAT GAC CAC TTA GGC | |
| CR6077335 | | TTG GGA GTT GTC TGA AGG C | chr3:93530292- 93530783 |
| | | GCT GGA TAC TGA ACA CAC G | |
| map16.1 | | CCT GTT CCT GTT CTA ATG TGC | chr3:93592781+ 93593080 |
| | | GAT AGT TTG GGA TTA GGG GAG | |
| S100A10 | exon 1 | GTT AGT CTG TTG GCA GAG TG | chr3:93554880+ 93555428 |
| | | AAG GAG AAC AGA ACA GCG G | |
| | exon 2 | CCA ATA CCT GAA GTG TCT GTT G | chr3:93560767+ 93561300 |
| | | CAC CTG CAG ATT GAG AGT TCT T | |
| | exon 3 | CCT TAT GCA CGT GGC TGT CA | chr3:93564119+ 93564794 |
| | | CCT GAG AGG AGA ATG CTT TCC C | |
| AI561653 | | GGA AAA ATA TTT CCC AGT TGC | chr3:93628599+ 93629848 |
| | | AGT TAT CTA TTA TTG TAT | |
| CJ058258 | | ACT CCT CAC AAC CAT CTG G | chr3:93640863+ 93641544 |
| | | TCT ACC CCA GAC ACA GAT TG | |
| tdpoz2 | | CCA ACA GGA AGA AGA ATA AAT A | chr3:93651548- 93652737 |
| | | TGT TGG TCA TAT CTG TCT AGC TC | |
| BB709516 | | CAA ACT GTG ACT CCC CAT C | chr3:93669214+ 93669833 |
| | | CCC TCT TCT GGT CTC CAT AG | |
| BY382353 | | GTT GGA TTT TAG AAG GCA GAA G | chr3:93680414+ 93680908 |
| | | TTC TTC CTT CAT CTT CTC TCC | |
| AK139746 | | CAG AAC TCG CTG GGA TAA GA | chr3:93920908+ 93921457 |
| | | AGA ATA AAA CAG GGC GTC CC | |
| tdpoz1 (tz1) | exon 2 | CTA TAA TTG TAT CTA ACA GGA TA | chr3:93873476- 93874475 |
| | | ATG TTG TGG AGG TCA GCC AG | |
| CN711625 | | TGG TTT GTC ATT TTG TGG AG | chr3:93781192+ 93782049 |
| | | TAA GGC TGC TCT GGG AAA CG | |

| | | | |
|---|---|---|---|
| | | TCA ACC TGT GTT TTT GGC | |
| | | AAA GGG TGC TGC TAA CTG GC | |
| BB712653 | | CAG AAG TCA GAG AAT CCT GC | chr3:93786587+ 93787164 |
| | | TTT CTT CAT CCC CTG TTT C | |
| map17.1 | | ACA AGC AAC AGA TTA CAC AGG | chr3:93802925+ 93803213 |
| | | TGA CTT CTT GGG AAT CTG TG | |
| tdpoz4 | | CAT AGA AGG GGC TAT TTA GGA G | chr3:93796258+ 93797685 |
| | | TTC CAG ATT CCA ATG CTG TC | |
| tdpoz3 | | TTG TAT CCA ACA GGA GGA AG | chr3:93825937+ 93827104 |
| | | GAT TCT AAG ACC ACA TTT GGC | |
| CN716290 | | CAG AGA AAC AGT GGG ACA GTC | chr3:93918928+ 93919523 |
| | | CAC AGA TTT TGT CAG CAG TTC | |
| AA415113 | | TGT GTA ACC ACT TGC TCC C | chr3:93942894+ 93943473 |
| | | CAG CAT CTA CTG AGG AAC ACC | |
| AU259387 | | TCC CAC CCT GGA TTT GAT GG | chr3:93950844+ 93951460 |
| | | CAA GGG GTA AAT GAC TGA GC | |
| tdpoz5 | | TGG ATA GAA TAG CCA AAA CG | chr3:93961373- 93961541 |
| | | TTA CTC TTC ATC CCG TTG G | |
| BB288298 | | CTT ACA AGG TTC AGG GTG G | chr3:93978562+ 93979034 |
| | | CGT GTC TTA TCC TTG TTA TTC G | |
| AA671366 | | GTG TGT GTT TTC CTG TCT CAC | chr3:93982925+ 93983680 |
| | | TGA GGG CAG TAA TGA CAA AC | |
| map18.1 | | GTG AGC ATA GTG GGA GCC TT | chr3:94010375+ 94010692 |
| | | GAG GTT TGG CAA CTT CCC AG | |
| AA684350 | | CAT TCC ACT TTC TGG GTC C | chr3:94010955+ 94011543 |
| | | GGT GGT TGA AAA ACT CTC CTG | |
| BB798131 | | GCA GTG GCA CAG ATG ATT ATC | chr3:94020895+ 94021532 |
| | | GGA GGA TGT TGA TTG ACC TG | |
| tdpoz1 | | AAC CTC CAA AAC TCC TCC AG | chr3:94048143- 94050144 |
| | | TTG ACC CAC AAA CAC TTC TC | |
| | | GAG AAA GAA AAA CAG TGT AGA GT | sequencing primers |
| | | TGC TGT CCC ACC AGT AGT AG | |
| gm9125 (AK033449) | | CTT CTA GTT CTT CTT CTC AGA C | chr3:94048656- 94050173 |
| | | TTC TCC TTT CCT CTG ACC AC | |
| | | CTT TTG AAA TCA GTC TGG ATT CC | chr3:94048067- 94048395 |
| | | AGA TGA GGT TAT ATA CAT GTG | |
| BY382353 (ii) | | GAG GTT GGA TTT TAG AAG GC | chr3:94058961+ 94059524 |
| | | ACC TGT GGC TCT TTC AGA C | |
| gm10697 | | TGT AAC CAA CAG GAT GAA GAG | chr3:94071505- 94072726 |
| | | CAC TGG TTA TTG CTA TTG GAA G | |
| AI467280 | | CGA GAA TGT CAA CAA AGT CC | chr3:94089722+ 94090444 |
| | | GCA TAA AAC CTT GTC TGT TGG | |
| AI605421 | | GAG AAA ATA GTC AGT GGC AAA C | chr3:94095337+ 94096031 |
| | | TGG ACA TTA GGT ATT GCC TG | |
| CN707578 | | GAC ACT GTT TTT TGC CAT TC | chr3:94113332+ 94114109 |
| | | TTG GTC GTC TAC CTT CCC TC | |
| BB802563 | | TGG CAC AGA TGA TTA GCA C | chr3:94133155+ |

| | | | |
|---|---|---|---|
| | | AGG ATG TTG ATT GAC CTG C | 94133789 |
| AA547104 | | TTT AGG TGT TTG AGT GAG GG | chr3:94157638+ 94158335 |
| | | CCT CAT CTC TCA TCT ATG TCG G | |
| gm10696 (tdpoz1) | | CAG ATC AAC TTA AAT CTC GA | chr3:94174318- 94176539 |
| | | CCG ATT TGT CCA GAA CTA TTC C | |
| | | GAG TGC TGA GAA AGA CTG ATG | |
| | | CGA GAA TGT CAA CAA AGT CC | |
| BB709439 | | GAC CAA CAT CTG TCT TCC AAC | chr3:94199472+ 94200172 |
| | | CAC AAA CAT CCA GAG CAG AG | |
| BB706344 | | TAT CAC GAA GCA GTG GCA C | chr3:94209295+ 94209874 |
| | | TCT TTT GGG GAA GTC CAG G | |
| BY683238 | | TTG TGA AGA GAA CCT GAA GC | chr3:94220809+ 94221349 |
| | | AGG TGG TTG TGT ATG AGG C | |
| map19.1 | | GTG TTG GTT CCC CTC AGT TC | chr3:94386613+ 94386865 |
| | | CAC CTG CTC AGT CCT TCA AG | |
| c2cd4d | exon 2 | CCA GGC GCA GCA CGC TGT CCT T | chr3:94363210- 94364035 |
| | | GAC TTC ACC TGT TCT TTT ACC TG | |
| Ling04 | exon 1 | CCC CAA GAG GGG CAG GGG CAC | chr3:94399207- 94399506 |
| | | CAA CAG GGT TGC TGC AGA CT | |
| | exon 2 | GTTGTGTGCCACACTTACTGGCA | chr3:94401699+ 94403684 |
| | | GAGACCTTAGTCCTACTTCTCC | |
| AK018134 | | AAA GAC ATC AAG GAA GGA GC | chr3:94432142+ 94432814 |
| | | AAT GGC AGA AGC AGC ATA G | |
| | | TGA GGT CTG TCA CTG CTA CG | chr3:94432673+ 94433382 |
| | | TTG CTG TTT CTT GGT GGG | |
| | | TCC CTC CCA CCA AGA AAC AG | chr3:94433360+ 94434053 |
| | | CTG AAA AAA GCA AAG CCA AC | |
| | | TTG GAG TGA TAG ACA CGG GC | chr3:94433760+ 94434522 |
| | | GGA GAA AAC CAA TGT GGA C | |
| | | CCA ATC TGA CCT TGA GTG C | chr3:94434091+ 94434748 |
| | | AGT CCT TTC CTC CAC GAA CC | |
| map20.1f | | AGC AGC ATG GGT AAG GAA AA | chr3:94693653+ 94693874 |
| | | GGT CCC TTT CTG TTC CTT AG | |
| 2310007A19 Rik | exon 4 | GTT CTT ACT AGC TCC CC | chr3:94472121+ 94472399 |
| | | GGT TAC CTC AAG AGA AAA TCC | |
| | exon 5 | ACT CCA CAA AAA GTC CCG GG | chr3:94466792+ 94467056 |
| | | GCT TCC TAA AGA GCA AGA TGG GGC | |
| | exon 6 | GTG ACA CAG AGA CCA ATT CCA | chr3:94465587+ 94465975 |
| | | GGA CAA AGC ATC CGA TGT T | |
| | exon 7 | CTC CAG GAG AAA AGG AGT CA | chr3:94464906+ 94465145 |
| | | TCC ACC TAA GGC CCC ACT CC | |

# APPENDIX III

## Complete list of SNPs discovered in *Matt*$^{ma/ma}$ mice compared to C57BL/6J animals

## APPENDIX III

Complete list of SNPs discovered in *Matt*$^{ma/ma}$ F6 mice compared to C57BL/6J animals. Identified between murine chromosome 3: 87,000,00 – 95,000,000 by homozygosity mapping.

| Position on chr3 | SNPs |
|---|---|
| 82,744,991 | HET A>G |
| 82,745,248 | HET T>C |
| 82,745,544 | HET A>G |
| 82,745,764 | HETA>G |
| 82,746,140 | HET G>T |
| 82,746,437 | HET G>A |
| 82,747,124 | HET T>C |
| 82,747,535 | HET C>T |
| 82,933,030 | HET G>A |
| 83,741,298 | HET A>G |
| 83,742,594 | HET A>G |
| 83,742,634 | HET C>T |
| 83,764,295 | HET T>C |
| 83,764,353 | HET T>C |
| 83,764,398 | HET INS 2 |
| 83,764,465 | HET T>C |
| 83,764,583 | HET G>A |
| 83,766,091 | HET INS 2 |
| 83,766,151 | HET T>C |
| 83,766,444 | HET A>C |
| 83,843,288 | HET G>C |
| 85,378,143 | HET DEL 4 |
| 85,719,618 | HET DEL 1 |
| 85,719,655 | HET A>G |
| 85,719,677 | HET C>A |
| 85,719,683 | HET G>A |
| 85,719,683 | HET del |
| 87,249,933 | HET C>A |
| 87,249,934 | HET G>A |
| 87,249,951 | HET G>T |
| 87,250,005 | HET G>T |
| 87,250,034 | HET C>T |
| 87,335,789 | DEL 17bp |
| 87,335,916 | HET G>A |
| 87,483,228 | HOM T>C |
| 87,483,926 | HOM G>A |
| 87,483,949 | HOM C>T |
| 87,495,710 | HOM T>C |
| 87,495,808 | HOM T>G |
| 87,495,828 | HOM T>G |
| 87,495,995 | HOM A>G |
| 87,496,248 | HOM A>G |
| 87,496,527 | HOM C>T |

| | |
|---|---|
| 87,496,588 | HOM G>A |
| 87,496,656 | HOM C>T |
| 87,497,092 | HOM A>C |
| 87,497,199 | HOM G>A |
| 87,540,755 | HOM A>G |
| 87,540,771 | HOM A>C |
| 87,540,969 | HOM T>A |
| 87,541,022 | HOM A>G |
| 87,541,031 | HOM G>A |
| 87,541,162 | HOM C>T |
| 87,544,483 | HOM DEL 8 |
| 87,544,594 | HOM A>C |
| 87,543,577 | HOM G>A |
| 87,546,241 | HOM C>G |
| 87,546,333 | HOM A>G |
| 87,546,514 | HOM G>A |
| 87,547,384 | HOM G>T |
| 87,547,732 | HOM A>G |
| 87,547,964 | HOM INS 1bp |
| 87,553,243 | HOM A>G |
| 87,553,531 | HOM G>A |
| 87,653,590 | HOM INS 8bp |
| 87,703,806 | HOM C>A |
| 87,715,662 | HOM G>A |
| 87,726,536 | HOM 2bp INS |
| 87,726,927 | HOM C>G |
| 87,830,377 | HOM A>G |
| 87,830,601 | HOM G>T |
| 87,830,703 | HOM DEL2 |
| 87,830,750 | HOM T>C |
| 87,830,778 | HOM C>T |
| 87,830,896 | HOM T>C |
| 87,830,902 | HOM A>G |
| 87,830,904 | HOM T>C |
| 87,830,932 | HOM C>T |
| 88,136,484 | HOM G>C |
| 88,508,770 | HET DEL 4 |
| 88,516,875 | HET INS 1 |
| 92,559,736 | HET C>A |
| 92,613,205 | HET A>C |
| 92,613,207 | HET A>C |
| 92,620,581 | HET T>G |
| 92,658,480 | HET C>T |
| 92,710,491 | HET T>C |
| 92,787,451 | HET C>G |
| 92,787,663 | HET C>T |
| 92,795,888 | HET ins 1 |
| 92,796,086 | HET A>G |
| 92,796,105 | HET G>A |
| 92,796,106 | HET A>C |
| 92,796,148 | HET C>T |
| 92,796,184 | HET A>T |
| 92,796,706 | HET T>G |

| | |
|---|---|
| 92,796,709 | HET T>G |
| 92,796,711 | HET T>C |
| 92,796,712 | HET G>A |
| 92,796,718 | HET C>G |
| 92,796,741 | HET A>C |
| 92,796,759 | HET delG:ins4 |
| 92,797,056 | HET G>A |
| 92,797,057 | HET C>T |
| 92,797,081 | HT G>C |
| 92,797,144 | HET ins3 |
| 92,797,151 | HET G>A |
| 92,806,140 | HET G>T |
| 92,806,148 | HET A>G |
| 92,806,225 | HET T>C |
| 92,806,393 | HET A>G |
| 92,806,475 | HET A>G |
| 92,831,634 | HET T>C |
| 92,845,913 | HET T>C |
| 92,846,091 | HET A>G |
| 92,904,126 | HET A>G |
| 92,904,346 | HET del 13 |
| 92,907,097 | HET A>G |
| 92,907,319 | HET G>A |
| 92,919,253 | HET ins/del 5 Gs |
| 92,930,010 | HET C>T |
| 92,930,041 | HET T>A |
| 92,940,197 | HET A>G |
| 92,940,280 | HET G>A |
| 92,948,823 | HET T>C |
| 92,951,438 | HET G>A |
| 92,951,451 | HET T>A |
| 92,951,587 | HET G>T |
| 92,951,632 | HET T>G |
| 92,951,714 | HET G>A |
| 92,952,141 | HET G>A |
| 92,952,150 | HET T>C |
| 92,952,427 | HET G>A |
| 92,952,934 | HET T>A |
| 92,952,939 | HET A>G |
| 92,977,775 | HET G>A |
| 93,004,013 | HET G>T |
| 93,004,386 | HET T>C |
| 93,066,337 | HET G>A |
| 93,066,427 | HET G>A |
| 93,077,695 | HET C>G |

# APPENDIX IV

## SNPs and InDels identified in *Matt*$^{ma/ma}$ by whole transcriptome sequencing

## Appendix IV

383 unique SNPs and InDels identified in F6 *Matt*$^{ma/ma}$ by whole transcriptome analysis (performed by Dr Christian Cole, University of Dundee). Regions between chromosome 3: 87,000,000 – 95,000-000 were analyzed using the Integrated Genome Browser (Nicol et al., 2009). 210 SNPs and InDels (orange) were located within the critical homozygous region chr3:87,336,555 – 88,392,125. Homozygous SNPs were investigated by checking all possible ORFs and transcripts in that region to shortlist those that caused a protein change (green). A single base change G>C was identified at 88,136,484 to cause a nonsense mutation, TYR>STOP (red).

Information legend:
DP = Raw read depth
VDB = Variant Distance Bias
AF1 = Max likelihood estimate of the first ALT allele frequency (assuming HWE)
AC1 = Max likelihood estimate of the first ALT allele count (no HWE assumption)
DP4 = # high-quality ref-forward bases, ref-reverse, alt-forward and alt-reverse bases
MQ = Root mean square mapping quality of covering reads
FQ = Phred probability of all samples being the same

| Position | Reference Base | Alternate Base | Quality | Information |
|---|---|---|---|---|
| 87052836 | C | A | 999 | DP=26;VDB=0.0400;AF1=1;AC1=6;DP4=0,0,13,13;MQ=42;FQ=-48.7 |
| 87055005 | C | T | 999 | DP=72;VDB=0.0329;AF1=1;AC1=6;DP4=0,0,69,2;MQ=27;FQ=-66.8 |
| 87055011 | C | T | 999 | DP=74;VDB=0.0372;AF1=1;AC1=6;DP4=0,0,67,2;MQ=27;FQ=-66.9 |
| 87055054 | G | A | 999 | DP=97;VDB=0.0417;AF1=1;AC1=6;DP4=0,0,62,32;MQ=34;FQ=-96.5 |
| 87055112 | G | A | 999 | DP=197;VDB=0.0477;AF1=1;AC1=6;DP4=0,0,87,91;MQ=39;FQ=-146 |
| 87055140 | T | G | 999 | DP=208;VDB=0.0525;AF1=1;AC1=6;DP4=0,0,107,78;MQ=41;FQ=-157 |
| 87055176 | T | C | 999 | DP=197;VDB=0.0522;AF1=1;AC1=6;DP4=0,0,107,73;MQ=40;FQ=-154 |
| 87055190 | T | G | 999 | DP=199;VDB=0.0522;AF1=1;AC1=6;DP4=0,0,116,64;MQ=40;FQ=-148 |
| 87056614 | T | C | 999 | DP=299;VDB=0.0471;AF1=1;AC1=6;DP4=0,0,146,126;MQ=40;FQ=-182 |
| 87056617 | A | C | 999 | DP=289;VDB=0.0436;AF1=1;AC1=6;DP4=0,0,149,124;MQ=40;FQ=-175 |
| 87061255 | A | G | 999 | DP=72;VDB=0.0029;AF1=1;AC1=6;DP4=0,0,35,35;MQ=39;FQ=-85 |

| | | | | |
|---|---|---|---|---|
| 87061311 | T | C | 999 | DP=184;VDB=0.0517;AF1=1;AC1=6;DP4=0,0,94,75;MQ=43;FQ=-148 |
| 87061405 | T | C | 999 | DP=164;VDB=0.0480;AF1=1;AC1=6;DP4=0,0,66,75;MQ=39;FQ=-118 |
| 87061417 | A | G | 999 | DP=159;VDB=0.0513;AF1=1;AC1=6;DP4=0,0,72,81;MQ=39;FQ=-112 |
| 87061441 | A | C | 999 | DP=151;VDB=0.0495;AF1=1;AC1=6;DP4=0,0,70,69;MQ=40;FQ=-100 |
| 87061495 | T | C | 999 | DP=83;VDB=0.0047;AF1=1;AC1=6;DP4=0,0,40,42;MQ=42;FQ=-74.2 |
| 87061510 | C | T | 999 | DP=70;VDB=0.0026;AF1=1;AC1=6;DP4=0,0,34,35;MQ=42;FQ=-69 |
| 87063385 | G | A | 999 | DP=104;VDB=0.0481;AF1=1;AC1=6;DP4=1,0,54,43;MQ=41;FQ=-88.9;PV4=1,1,1,1 |
| 87063487 | C | T | 151 | DP=105;VDB=0.0528;AF1=0.8395;AC1=5;DP4=0,1,9,8;MQ=43;FQ=-17;PV4=1,0.00098,0.15,1 |
| 87067404 | C | A | 999 | DP=85;VDB=0.0036;AF1=1;AC1=6;DP4=0,0,4,77;MQ=37;FQ=-75.9 |
| 87067464 | A | G | 999 | DP=50;VDB=0.0225;AF1=1;AC1=6;DP4=0,1,1,46;MQ=36;FQ=-38;PV4=1,1,0.00096,0.16 |
| 87329777 | C | T | 999 | DP=40;VDB=0.0350;AF1=1;AC1=6;DP4=0,0,24,14;MQ=40;FQ=-65.6 |
| 87330260 | T | C | 999 | DP=22;VDB=0.0414;AF1=1;AC1=6;DP4=0,0,11,10;MQ=47;FQ=-42.4 |
| 87333542 | C | T | 999 | DP=47;VDB=0.0425;AF1=1;AC1=6;DP4=0,0,21,24;MQ=49;FQ=-68.2 |
| 87333810 | CTTT | CTT | 999 | INDEL;DP=129;VDB=0.0510;AF1=1;AC1=6;DP4=0,0,35,82;MQ=49;FQ=-126 |
| 87333864 | C | T | 999 | DP=88;VDB=0.0508;AF1=1;AC1=6;DP4=0,0,15,66;MQ=48;FQ=-96.9 |
| 87333939 | CGGGG | CGGG | 23.8 | INDEL;DP=21;VDB=0.0309;AF1=0.7538;AC1=5;DP4=0,1,2,1;MQ=41;FQ=-20.3;PV4=1,0.33,0.21,1 |
| 87334829 | G | T | 999 | DP=27;VDB=0.0452;AF1=1;AC1=6;DP4=0,0,8,16;MQ=42;FQ=-52.5 |
| 87334852 | A | C | 999 | DP=20;VDB=0.0422;AF1=1;AC1=6;DP4=0,0,6,13;MQ=45;FQ=-42.9 |
| 87335336 | T | C | 13.6 | DP=25;VDB=0.0386;AF1=0.1751;AC1=1;DP4=13,7,2,1;MQ=43;FQ=14.5;PV4=1,0.00013,0.022,1 |
| 87335361 | A | G | 999 | DP=23;VDB=0.0411;AF1=1;AC1=6;DP4=0,0,9,12;MQ=36;FQ=-42.8 |
| 87335365 | A | G | 999 | DP=23;VDB=0.0374;AF1=1;AC1=6;DP4=0,0,9,14;MQ=37;FQ=-42.8 |
| 87335778 | C | T | 999 | DP=21;VDB=0.0475;AF1=1;AC1=6;DP4=0,0,12,9;MQ=50;FQ=-48.7 |
| 87335915 | G | A | 999 | DP=20;VDB=0.0457;AF1=1;AC1=6;DP4=0,0,7,13;MQ=45;FQ=-47.6 |
| 87336555 | A | C | 999 | DP=32;VDB=0.0396;AF1=1;AC1=6;DP4=0,0,19,10;MQ=47;FQ=-56.6 |
| 87338156 | G | A | 999 | DP=25;VDB=0.0410;AF1=1;AC1=6;DP4=0,0,14,8;MQ=46;FQ=-50.2 |
| 87338161 | CA | CAA | 999 | INDEL;DP=26;VDB=0.0443;AF1=1;AC1=6;DP4=0,0,17,9;MQ=43;FQ=-58.9 |
| 87340061 | C | T | 999 | DP=653;VDB=0.0421;AF1=1;AC1=6;DP4=1,0,242,357;MQ=47;FQ=- |

| | | | | |
|---|---|---|---|---|
| | | | | 288;PV4=0.41,0.00014,1,1 |
| 87340262 | G | T | 999 | DP=711;VDB=0.0528;AF1=1;AC1=6;DP4=0,0,411,261;MQ=48;FQ=-288 |
| 87340393 | T | C | 999 | DP=663;VDB=0.0503;AF1=1;AC1=6;DP4=0,0,248,378;MQ=48;FQ=-288 |
| 87340851 | T | C | 999 | DP=679;VDB=0.0459;AF1=1;AC1=6;DP4=1,0,346,282;MQ=44;FQ=-288;PV4=1,1,0.013,1 |
| 87341147 | T | C | 999 | DP=686;VDB=0.0509;AF1=1;AC1=6;DP4=0,1,358,145;MQ=39;FQ=-288;PV4=0.29,1,1,1 |
| 87341169 | C | T | 999 | DP=832;VDB=0.0494;AF1=1;AC1=6;DP4=0,1,343,304;MQ=40;FQ=-288;PV4=0.47,0.39,1,1 |
| 87341384 | A | G | 999 | DP=413;VDB=0.0519;AF1=1;AC1=6;DP4=0,0,155,195;MQ=48;FQ=-286 |
| 87341513 | ATTTTCTTTTTCTTTTTCTTTC | ATTTTCTTTTTCTTTTC | 999 | INDEL;DP=356;VDB=0.0421;AF1=1;AC1=6;DP4=0,0,254,40;MQ=34;FQ=-277 |
| 87341855 | T | A | 999 | DP=1424;VDB=0.0392;AF1=1;AC1=6;DP4=0,0,845,510;MQ=44;FQ=-288 |
| 87342038 | A | G | 999 | DP=1021;VDB=0.0523;AF1=1;AC1=6;DP4=0,0,377,581;MQ=44;FQ=-288 |
| 87342382 | G | A | 999 | DP=895;VDB=0.0526;AF1=1;AC1=6;DP4=1,1,234,332;MQ=49;FQ=-207;PV4=1,1.8e-09,1,0.38 |
| 87342549 | G | A | 999 | DP=1417;VDB=0.0517;AF1=1;AC1=6;DP4=1,1,820,429;MQ=49;FQ=-288;PV4=1,7.4e-07,0.31,1 |
| 87342794 | A | G | 999 | DP=922;VDB=0.0528;AF1=1;AC1=6;DP4=0,0,265,500;MQ=49;FQ=-288 |
| 87342954 | C | T | 999 | DP=1094;VDB=0.0522;AF1=1;AC1=6;DP4=1,0,272,322;MQ=49;FQ=-224;PV4=0.46,8.3e-16,0.43,1 |
| 87343374 | G | A | 999 | DP=620;VDB=0.0517;AF1=1;AC1=6;DP4=0,0,328,256;MQ=48;FQ=-288 |
| 87343505 | TAAAAAAAAAAAA | TAAAAAAAAAAAA,TAAAAAAAAAAAAA,TAAAAAAAAA | 186 | INDEL;DP=619;VDB=0.0461;AF1=1;AC1=6;DP4=3,11,196,252;MQ=50;FQ=-122;PV4=0.11,1,0.35,0.22 |
| 87422049 | G | T | 999 | DP=91;VDB=0.0494;AF1=1;AC1=6;DP4=0,0,42,41;MQ=47;FQ=-98.2 |
| 87422125 | TTTTTGTTTTGTTTTGTTTTGTTTT | TTTTTGTTTTGTTTTGTTTTTGTTTTGTTTT | 999 | INDEL;DP=70;VDB=0.0466;AF1=1;AC1=6;DP4=0,0,29,23;MQ=33;FQ=-79.5 |
| 87422209 | ATTTTTTTTTTT | ATTTTTTTTTTTTTT,ATTTTTTTTTTTTTTT,ATTTTTTTTTTTTT | 999 | INDEL;DP=79;VDB=0.0406;AF1=1;AC1=6;DP4=0,0,21,48;MQ=42;FQ=-78.4 |
| 87422423 | CT | CTTTTT,CTTTT | 110 | INDEL;DP=53;VDB=0.0375;AF1=1;AC1=6;DP4=0,0,5,2;MQ=29;FQ=-39.3 |

| | | | | |
|---|---|---|---|---|
| 87422441 | TTTCTT | TTTCTTCTT | 999 | INDEL;DP=77;VDB=0.0457;AF1=1;AC1=6;DP4=0,0,23,7;MQ=29;FQ=-63.7 |
| 87422503 | C | T | 999 | DP=164;VDB=0.0281;AF1=1;AC1=6;DP4=0,0,110,43;MQ=45;FQ=-145 |
| 87491844 | T | C | 999 | DP=259;VDB=0.0519;AF1=1;AC1=6;DP4=0,1,110,132;MQ=36;FQ=-200;PV4=1,0.4,1,0.18 |
| 87501985 | AGG | AG | 30.5 | INDEL;DP=25;VDB=0.0127;AF1=0.6228;AC1=4;DP4=1,0,2,1;MQ=41;FQ=3.23;PV4=1,0.013,0,0.00066 |
| 87521343 | C | A | 4.95 | DP=25;VDB=0.0053;AF1=0.1679;AC1=1;DP4=4,15,2,1;MQ=40;FQ=5.51;PV4=0.17,1.1e-07,0.16,0.47 |
| 87526274 | TG | TGG | 76.7 | INDEL;DP=26;VDB=0.0147;AF1=0.3726;AC1=2;DP4=4,4,2,3;MQ=42;FQ=78.8;PV4=1,1,0.046,0.00082 |
| 87530398 | T | G,A | 999 | DP=191;VDB=0.0076;AF1=1;AC1=6;DP4=0,0,77,110;MQ=37;FQ=-194 |
| 87531381 | A | G | 999 | DP=188;VDB=0.0076;AF1=1;AC1=6;DP4=0,0,89,77;MQ=36;FQ=-163 |
| 87531890 | C | T,G | 999 | DP=70;VDB=0.0004;AF1=1;AC1=6;DP4=0,0,9,25;MQ=36;FQ=-59.1 |
| 87532452 | T | C | 999 | DP=104;VDB=0.0007;AF1=1;AC1=6;DP4=0,0,34,64;MQ=38;FQ=-96.8 |
| 87533405 | A | G | 117 | DP=54;VDB=0.0000;AF1=0.9323;AC1=5;DP4=0,1,9,8;MQ=39;FQ=-27.1;PV4=1,2.7e-10,0.16,0.00073 |
| 87533576 | TA | TAA | 23.7 | INDEL;DP=58;VDB=0.0018;AF1=0.6263;AC1=4;DP4=1,4,6,0;MQ=38;FQ=27.9;PV4=0.015,0.00018,0.074,0.042 |
| 87536587 | G | A | 999 | DP=73;VDB=0.0114;AF1=1;AC1=6;DP4=0,0,34,32;MQ=42;FQ=-81.9 |
| 87538170 | G | A | 999 | DP=22;VDB=0.0246;AF1=1;AC1=6;DP4=0,0,12,6;MQ=48;FQ=-41.9 |
| 87538873 | T | C | 999 | DP=21;VDB=0.0433;AF1=1;AC1=6;DP4=0,0,10,10;MQ=45;FQ=-47.6 |
| 87541061 | G | A | 999 | DP=730;VDB=0.0520;AF1=1;AC1=6;DP4=0,0,434,243;MQ=39;FQ=-288 |
| 87541097 | A | G | 999 | DP=708;VDB=0.0520;AF1=1;AC1=6;DP4=0,0,384,292;MQ=40;FQ=-288 |
| 87541294 | A | G | 999 | DP=760;VDB=0.0528;AF1=1;AC1=6;DP4=1,0,257,447;MQ=44;FQ=-288;PV4=0.37,0.059,0.021,0.46 |
| 87541653 | A | G | 999 | DP=638;VDB=0.0517;AF1=1;AC1=6;DP4=0,1,327,276;MQ=44;FQ=-288;PV4=0.46,0.32,0.033,0.33 |
| 87551922 | G | A | 999 | DP=30;VDB=0.0459;AF1=1;AC1=6;DP4=0,0,13,14;MQ=49;FQ=-45.9 |
| 87552005 | A | G | 999 | DP=43;VDB=0.0348;AF1=1;AC1=6;DP4=0,0,27,12;MQ=38;FQ=-52 |
| 87552010 | G | A | 999 | DP=44;VDB=0.0297;AF1=1;AC1=6;DP4=0,0,27,14;MQ=38;FQ=-52 |
| 87552429 | G | A | 999 | DP=29;VDB=0.0503;AF1=1;AC1=6;DP4=0,0,15,13;MQ=49;FQ=-51.9 |
| 87552560 | C | T | 999 | DP=38;VDB=0.0314;AF1=1;AC1=6;DP4=0,0,20,16;MQ=46;FQ=-57.8 |
| 87552569 | T | C | 999 | DP=45;VDB=0.0419;AF1=1;AC1=6;DP4=0,0,22,19;MQ=45;FQ=-57.8 |
| 87552622 | A | T | 999 | DP=51;VDB=0.0424;AF1=1;AC1=6;DP4=0,0,19,31;MQ=48;FQ=-61 |

| | | | | |
|---|---|---|---|---|
| 87552820 | T | G | 999 | DP=49;VDB=0.0467;AF1=1;AC1=6;DP4=0,0,14,34;MQ=46;FQ=-69.8 |
| 87552836 | T | C | 999 | DP=49;VDB=0.0456;AF1=1;AC1=6;DP4=0,0,15,30;MQ=45;FQ=-63.9 |
| 87553242 | A | G | 999 | DP=136;VDB=0.0530;AF1=1;AC1=6;DP4=0,1,66,57;MQ=49;FQ=-124;PV4=0.47,0.0075,0.41,0.15 |
| 87553530 | G | A | 999 | DP=191;VDB=0.0397;AF1=1;AC1=6;DP4=0,0,78,102;MQ=48;FQ=-166 |
| 87553820 | T | A | 106 | DP=118;VDB=0.0509;AF1=1;AC1=6;DP4=0,0,6,4;MQ=50;FQ=-36 |
| 87553902 | G | T,C | 999 | DP=260;VDB=0.0523;AF1=1;AC1=6;DP4=0,0,157,78;MQ=49;FQ=-226 |
| 87554032 | G | A | 999 | DP=504;VDB=0.0528;AF1=1;AC1=6;DP4=0,0,269,160;MQ=47;FQ=-248 |
| 87555066 | A | G | 999 | DP=235;VDB=0.0070;AF1=1;AC1=6;DP4=0,0,102,114;MQ=38;FQ=-185 |
| 87555168 | G | A | 999 | DP=228;VDB=0.0480;AF1=1;AC1=6;DP4=0,0,69,141;MQ=38;FQ=-182 |
| 87555381 | C | T | 999 | DP=37;VDB=0.0001;AF1=1;AC1=6;DP4=0,0,10,26;MQ=36;FQ=-46 |
| 87555576 | T | G | 999 | DP=27;VDB=0.0001;AF1=1;AC1=6;DP4=0,0,12,6;MQ=35;FQ=-46.5 |
| 87555619 | A | G | 999 | DP=145;VDB=0.0221;AF1=1;AC1=6;DP4=0,0,46,34;MQ=35;FQ=-96.7 |
| 87555620 | C | T | 999 | DP=152;VDB=0.0235;AF1=1;AC1=6;DP4=0,0,41,34;MQ=35;FQ=-93.5 |
| 87555666 | G | A | 999 | DP=116;VDB=0.0045;AF1=1;AC1=6;DP4=0,0,61,50;MQ=36;FQ=-120 |
| 87555696 | CT | CTT | 999 | INDEL;DP=30;VDB=0.0044;AF1=1;AC1=6;DP4=0,0,16,9;MQ=35;FQ=-53 |
| 87555788 | T | C | 999 | DP=80;VDB=0.0011;AF1=1;AC1=6;DP4=0,0,41,27;MQ=36;FQ=-81.7 |
| 87556028 | C | T | 7.9 | DP=27;VDB=0.0000;AF1=0.483;G3=5.005e-08,1,1.203e-37;HWE=0.0467;AC1=3;DP4=6,1,2,4;MQ=46;FQ=10.3;PV4=0.1,1.3e-08,0.12,0.00011 |
| 87556195 | A | G | 999 | DP=193;VDB=0.0281;AF1=1;AC1=6;DP4=0,1,84,106;MQ=39;FQ=-169;PV4=1,0.26,0.069,0.23 |
| 87558498 | G | A | 999 | DP=317;VDB=0.0500;AF1=1;AC1=6;DP4=0,1,179,129;MQ=36;FQ=-287;PV4=0.42,1,0.0077,1 |
| 87559852 | G | A | 999 | DP=154;VDB=0.0017;AF1=1;AC1=6;DP4=0,1,78,72;MQ=39;FQ=-136;PV4=0.48,0.47,1,0.086 |
| 87560037 | C | T | 999 | DP=103;VDB=0.0001;AF1=1;AC1=6;DP4=1,0,25,53;MQ=38;FQ=-95.7;PV4=0.33,0.0003,1,1 |
| 87560049 | G | A | 999 | DP=179;VDB=0.0010;AF1=1;AC1=6;DP4=0,0,58,104;MQ=40;FQ=-165 |
| 87560581 | A | G | 999 | DP=251;VDB=0.0520;AF1=1;AC1=6;DP4=0,0,74,152;MQ=39;FQ=-185 |
| 87561962 | G | A | 999 | DP=307;VDB=0.0507;AF1=1;AC1=6;DP4=1,0,201,78;MQ=42;FQ=-221;PV4=1,1.8e-08,0.18,0.19 |
| 87562009 | T | C | 999 | DP=347;VDB=0.0531;AF1=1;AC1=6;DP4=0,1,152,155;MQ=44;FQ=-231;PV4=1,3.5e-05,0.22,1 |

| | | | | |
|---|---|---|---|---|
| 87566201 | ATT | ATTT | 115 | INDEL;DP=20;VDB=0.0069;AF1=0.5003;G3=4.645e-09,1,1.85e-08;HWE=0.0418;AC1=3;DP4=1,8,2,4;MQ=37;FQ=118;PV4=0.53,1,1,0.36 |
| 87671334 | T | TG | 999 | INDEL;DP=128;VDB=0.0002;AF1=1;AC1=6;DP4=1,0,105,16;MQ=37;FQ=-125;PV4=1,1,0.0064,0.00017 |
| 87673799 | C | T | 999 | DP=1501;VDB=0.0510;AF1=1;AC1=6;DP4=3,0,745,660;MQ=49;FQ=-288;PV4=0.25,6.3e-15,0.3,1 |
| 87710637 | G | A | 8.42 | DP=69;VDB=0.0001;AF1=0.5696;AC1=4;DP4=3,0,4,1;MQ=43;FQ=11.6;PV4=1,7.2e-06,0.026,0.006 |
| 87719017 | C | T | 999 | DP=4523;VDB=0.0484;AF1=1;AC1=6;DP4=0,1,1685,2686;MQ=41;FQ=-288;PV4=1,1,0.24,0.16 |
| 87720039 | C | T | 999 | DP=147;VDB=0.0014;AF1=0.8354;AC1=5;DP4=0,7,2,105;MQ=47;FQ=-11.3;PV4=1,2.9e-09,0.098,1 |
| 87723551 | T | A | 999 | DP=613;VDB=0.0467;AF1=1;AC1=6;DP4=0,0,490,91;MQ=43;FQ=-288 |
| 87723928 | ACCCCCCCCCCCC | ACCCCCCCCCCC,ACCCCCCCCCCC,ACCCCCCCCC | 70.2 | INDEL;DP=74;VDB=0.0421;AF1=1;AC1=6;DP4=0,0,18,18;MQ=47;FQ=-46.4 |
| 87723979 | A | G | 999 | DP=124;VDB=0.0481;AF1=1;AC1=6;DP4=0,0,56,46;MQ=44;FQ=-120 |
| 87724007 | C | A | 999 | DP=140;VDB=0.0482;AF1=1;AC1=6;DP4=0,0,83,53;MQ=43;FQ=-157 |
| 87725443 | A | G | 29.5 | DP=92;VDB=0.0015;AF1=0.33;AC1=2;DP4=1,82,0,8;MQ=49;FQ=31.2;PV4=1,1,1.7e-17,1 |
| 87725448 | A | G | 13.1 | DP=92;VDB=0.0004;AF1=0.468;AC1=3;DP4=1,72,0,19;MQ=49;FQ=15.6;PV4=1,5.8e-29,0.084,0.17 |
| 87725564 | T | C | 262 | DP=33;VDB=0.0296;AF1=1;AC1=6;DP4=0,0,16,4;MQ=42;FQ=-47.6 |
| 87725609 | A | G | 7.71 | DP=100;VDB=0.0328;AF1=0.3231;AC1=2;DP4=65,14,16,3;MQ=36;FQ=9.12;PV4=1,2.4e-21,1,1 |
| 87726828 | ATTTTTTTTTTTTTT | ATTTTTTTTTTTT,ATTTTTTTTTTTT | 246 | INDEL;DP=60;VDB=0.0510;AF1=1;AC1=6;DP4=0,0,17,18;MQ=45;FQ=-66.6 |
| 87726926 | C | G | 999 | DP=1482;VDB=0.0449;AF1=1;AC1=6;DP4=1,2,689,659;MQ=40;FQ=-288;PV4=0.62,8.5e-11,1,1 |
| 87728507 | C | G | 999 | DP=95;VDB=0.0522;AF1=1;AC1=6;DP4=0,0,51,39;MQ=49;FQ=-106 |
| 87729141 | T | C | 999 | DP=234;VDB=0.0526;AF1=1;AC1=6;DP4=0,0,90,133;MQ=48;FQ=-215 |
| 87729935 | G | T | 999 | DP=21;VDB=0.0476;AF1=1;AC1=6;DP4=0,0,12,8;MQ=48;FQ=-48.4 |
| 87731484 | C | A | 5.24 | DP=21;VDB=0.0007;AF1=0.3372;AC1=2;DP4=4,3,3,1;MQ=46;FQ=6.52;PV4=1,6.6e-06,0.002,0.0037 |
| 87732377 | A | G | 249 | DP=25;VDB=0.0493;AF1=1;AC1=6;DP4=0,0,14,9;MQ=49;FQ=-45 |
| 87732464 | A | T | 228 | DP=21;VDB=0.0253;AF1=1;AC1=6;DP4=0,0,8,13;MQ=50;FQ=-47.8 |

| | | | | |
|---|---|---|---|---|
| 87733197 | T | G | 999 | DP=60;VDB=0.0522;AF1=1;AC1=6;DP4=0,0,28,12;MQ=49;FQ=-63 |
| 87734692 | T | C | 999 | DP=36;VDB=0.0487;AF1=1;AC1=6;DP4=0,0,15,18;MQ=48;FQ=-57.4 |
| 87735221 | T | C | 999 | DP=58;VDB=0.0356;AF1=1;AC1=6;DP4=0,0,45,11;MQ=44;FQ=-72.8 |
| 87735633 | AAAGAAGAA | AAAGAAGAAGAA | 999 | INDEL;DP=288;VDB=0.0510;AF1=1;AC1=6;DP4=0,0,81,143;MQ=30;FQ=-211 |
| 87735674 | A | G | 999 | DP=302;VDB=0.0471;AF1=1;AC1=6;DP4=0,0,162,129;MQ=36;FQ=-251 |
| 87736062 | A | G | 999 | DP=1435;VDB=0.0504;AF1=1;AC1=6;DP4=1,0,761,544;MQ=41;FQ=-288;PV4=1,1,1,0.16 |
| 87736074 | T | C | 999 | DP=1471;VDB=0.0523;AF1=1;AC1=6;DP4=1,0,748,560;MQ=40;FQ=-288;PV4=1,1,1,0.37 |
| 87743985 | A | G | 999 | DP=499;VDB=0.0484;AF1=1;AC1=6;DP4=0,0,245,200;MQ=46;FQ=-288 |
| 87744075 | C | G | 999 | DP=624;VDB=0.0436;AF1=1;AC1=6;DP4=0,0,332,248;MQ=32;FQ=-288 |
| 87756996 | C | T | 999 | DP=1919;VDB=0.0510;AF1=1;AC1=6;DP4=1,1,895,904;MQ=43;FQ=-288;PV4=1,0.00021,1,0.13 |
| 87757110 | A | C | 999 | DP=1301;VDB=0.0504;AF1=1;AC1=6;DP4=1,0,329,914;MQ=46;FQ=-288;PV4=0.27,1.4e-07,0.26,0.37 |
| 87757183 | C | T | 999 | DP=724;VDB=0.0457;AF1=1;AC1=6;DP4=0,2,83,601;MQ=44;FQ=-288;PV4=1,0.38,0.04,1 |
| 87775838 | G | A | 999 | DP=180;VDB=0.0461;AF1=1;AC1=6;DP4=0,0,69,103;MQ=44;FQ=-188 |
| 87778767 | A | G | 999 | DP=25;VDB=0.0011;AF1=1;AC1=6;DP4=0,0,17,7;MQ=38;FQ=-52.5 |
| 87779341 | A | T | 999 | DP=29;VDB=0.0001;AF1=0.8314;AC1=5;DP4=0,2,7,16;MQ=39;FQ=-4.7;PV4=1,0.31,0.022,0.03 |
| 87779419 | A | G | 999 | DP=211;VDB=0.0503;AF1=1;AC1=6;DP4=0,0,124,82;MQ=44;FQ=-172 |
| 87779590 | C | T | 999 | DP=199;VDB=0.0526;AF1=1;AC1=6;DP4=0,0,94,77;MQ=48;FQ=-160 |
| 87779650 | A | G | 999 | DP=220;VDB=0.0514;AF1=1;AC1=6;DP4=0,0,109,91;MQ=48;FQ=-182 |
| 87779729 | A | T | 999 | DP=193;VDB=0.0519;AF1=1;AC1=6;DP4=0,0,78,89;MQ=47;FQ=-182 |
| 87779961 | T | A | 999 | DP=137;VDB=0.0447;AF1=1;AC1=6;DP4=0,0,55,75;MQ=47;FQ=-133 |
| 87779987 | G | A | 999 | DP=147;VDB=0.0478;AF1=1;AC1=6;DP4=0,0,61,82;MQ=47;FQ=-145 |
| 87780029 | G | A | 999 | DP=166;VDB=0.0520;AF1=1;AC1=6;DP4=0,0,73,88;MQ=48;FQ=-175 |
| 87780066 | C | T | 999 | DP=133;VDB=0.0528;AF1=1;AC1=6;DP4=0,0,52,65;MQ=48;FQ=-129 |
| 87780117 | T | C | 999 | DP=130;VDB=0.0398;AF1=1;AC1=6;DP4=0,0,63,62;MQ=49;FQ=-115 |
| 87780274 | G | A | 999 | DP=181;VDB=0.0517;AF1=1;AC1=6;DP4=0,0,82,90;MQ=48;FQ=-172 |
| 87780298 | A | G | 999 | DP=194;VDB=0.0521;AF1=1;AC1=6;DP4=0,0,68,102;MQ=48;FQ=-184 |
| 87780376 | G | A | 999 | DP=142;VDB=0.0447;AF1=1;AC1=6;DP4=0,0,74,64;MQ=44;FQ=-142 |

| | | | | |
|---|---|---|---|---|
| 87780532 | G | A | 999 | DP=110;VDB=0.0425;AF1=1;AC1=6;DP4=1,0,75,31;MQ=36;FQ=-112;PV4=1,1,0.15,1 |
| 87780559 | A | T | 146 | DP=74;VDB=0.0509;AF1=0.679;AC1=4;DP4=16,1,40,8;MQ=15;FQ=-5.61;PV4=0.43,1.2e-20,0.33,1 |
| 87780622 | C | A | 999 | DP=128;VDB=0.0345;AF1=1;AC1=6;DP4=0,0,72,44;MQ=32;FQ=-136 |
| 87780909 | T | C | 999 | DP=133;VDB=0.0429;AF1=1;AC1=6;DP4=0,0,15,12;MQ=41;FQ=-53.8 |
| 87780910 | G | A | 999 | DP=133;VDB=0.0429;AF1=1;AC1=6;DP4=0,0,16,13;MQ=41;FQ=-53.9 |
| 87780921 | A | G | 999 | DP=131;VDB=0.0476;AF1=1;AC1=6;DP4=0,0,80,48;MQ=45;FQ=-124 |
| 87781057 | A | C | 999 | DP=196;VDB=0.0527;AF1=1;AC1=6;DP4=0,0,96,79;MQ=48;FQ=-148 |
| 87781379 | A | G | 999 | DP=244;VDB=0.0482;AF1=1;AC1=6;DP4=1,2,113,98;MQ=49;FQ=-151;PV4=0.6,9.7e-15,0.32,1 |
| 87781933 | G | A | 999 | DP=307;VDB=0.0484;AF1=1;AC1=6;DP4=0,1,106,183;MQ=47;FQ=-236;PV4=1,0.00057,0.28,1 |
| 87782080 | AG | AGTGG | 999 | INDEL;DP=233;VDB=0.0449;AF1=1;AC1=6;DP4=0,0,85,139;MQ=46;FQ=-250 |
| 87782338 | A | G | 999 | DP=402;VDB=0.0510;AF1=1;AC1=6;DP4=0,1,141,230;MQ=48;FQ=-287;PV4=1,4e-08,0.29,0.18 |
| 87782401 | G | A | 999 | DP=460;VDB=0.0461;AF1=1;AC1=6;DP4=0,0,230,215;MQ=48;FQ=-288 |
| 87782567 | G | A | 999 | DP=344;VDB=0.0513;AF1=1;AC1=6;DP4=0,1,92,233;MQ=48;FQ=-211;PV4=1,0.00085,0.29,0.4 |
| 87782585 | G | A | 999 | DP=350;VDB=0.0461;AF1=1;AC1=6;DP4=0,1,127,201;MQ=48;FQ=-219;PV4=1,1.5e-08,0.31,0.2 |
| 87783129 | C | T | 999 | DP=489;VDB=0.0419;AF1=1;AC1=6;DP4=0,0,258,216;MQ=49;FQ=-288 |
| 87783286 | C | A | 999 | DP=267;VDB=0.0517;AF1=1;AC1=6;DP4=1,0,60,179;MQ=49;FQ=-212;PV4=0.25,1.2e-09,1,0.18 |
| 87783407 | A | G | 999 | DP=325;VDB=0.0444;AF1=1;AC1=6;DP4=0,0,240,66;MQ=47;FQ=-250 |
| 87783412 | G | A | 999 | DP=330;VDB=0.0444;AF1=1;AC1=6;DP4=1,0,247,75;MQ=47;FQ=-288;PV4=1,0.00043,1,1 |
| 87783616 | T | A | 999 | DP=430;VDB=0.0495;AF1=1;AC1=6;DP4=0,0,214,176;MQ=48;FQ=-288 |
| 87783819 | G | A | 999 | DP=370;VDB=0.0524;AF1=1;AC1=6;DP4=3,1,179,149;MQ=48;FQ=-283;PV4=0.63,0.00092,1,1 |
| 87783989 | TGGGGGG | TGGGGGGGG | 999 | INDEL;DP=184;VDB=0.0481;AF1=1;AC1=6;DP4=0,0,88,76;MQ=46;FQ=-176 |
| 87784005 | C | T | 999 | DP=258;VDB=0.0490;AF1=1;AC1=6;DP4=0,0,132,106;MQ=47;FQ=-216 |
| 87784180 | T | C | 999 | DP=472;VDB=0.0529;AF1=1;AC1=6;DP4=0,0,151,289;MQ=48;FQ=-277 |
| 87784204 | A | G | 999 | DP=476;VDB=0.0517;AF1=1;AC1=6;DP4=0,0,137,302;MQ=48;FQ=-288 |

| | | | | |
|---|---|---|---|---|
| 87858381 | AT | ATT | 999 | INDEL;DP=29;VDB=0.0043;AF1=0.7856;AC1=5;DP4=0,2,7,12;MQ=33;FQ=-20.1;PV4=0.53,1,1,0.0001 |
| 87858954 | A | G | 999 | DP=332;VDB=0.0421;AF1=1;AC1=6;DP4=0,0,138,182;MQ=40;FQ=-288 |
| 87859337 | C | T | 999 | DP=594;VDB=0.0500;AF1=1;AC1=6;DP4=0,0,296,273;MQ=47;FQ=-288 |
| 87860731 | G | A | 999 | DP=27;VDB=0.0462;AF1=1;AC1=6;DP4=0,0,11,14;MQ=48;FQ=-52.8 |
| 87860892 | A | AT | 14.8 | INDEL;DP=36;VDB=0.0470;AF1=0.1667;AC1=1;DP4=22,8,2,1;MQ=47;FQ=15.6;PV4=1,0.075,0.17,0.00073 |
| 87862299 | T | C | 999 | DP=902;VDB=0.0345;AF1=1;AC1=6;DP4=0,0,98,103;MQ=41;FQ=-212 |
| 87862300 | G | A | 999 | DP=907;VDB=0.0361;AF1=1;AC1=6;DP4=0,1,94,101;MQ=41;FQ=-199;PV4=1,0.28,1,1 |
| 87905349 | G | A | 21.2 | DP=36;VDB=0.0000;AF1=0.58;AC1=4;DP4=3,5,0,8;MQ=44;FQ=24.8;PV4=0.2,1.4e-11,0,0.00012 |
| 87908295 | GT | GTT | 5.16 | INDEL;DP=22;VDB=0.0213;AF1=0.4253;AC1=3;DP4=2,3,3,0;MQ=35;FQ=6.84;PV4=0.2,0.048,0.13,0.027 |
| 87912435 | G | GA | 999 | INDEL;DP=25;VDB=0.0001;AF1=1;AC1=6;DP4=0,0,15,5;MQ=41;FQ=-52.9 |
| 87913417 | AGG | AG | 18.2 | INDEL;DP=55;VDB=0.0024;AF1=0.1846;AC1=1;DP4=1,7,2,1;MQ=48;FQ=19.1;PV4=0.15,0.031,0.0028,0.067 |
| 87919821 | C | T | 8.83 | DP=46;VDB=0.0043;AF1=0.3904;AC1=2;DP4=1,1,1,1;MQ=50;FQ=10.6;PV4=1,0.045,0.21,0.23 |
| 87950762 | T | A | 27.3 | DP=21;VDB=0.0340;AF1=0.6661;AC1=4;DP4=1,1,1,2;MQ=49;FQ=16;PV4=1,0.016,0.027,1 |
| 87959598 | A | T | 10.1 | DP=63;VDB=0.0000;AF1=0.2746;AC1=2;DP4=13,8,6,3;MQ=47;FQ=11.3;PV4=1,1.8e-14,0.007,0.00031 |
| 87971431 | C | CAA | 123 | INDEL;DP=27;VDB=0.0299;AF1=0.4967;G3=1.6e-08,1,3.988e-26;HWE=0.0424;AC1=3;DP4=2,6,0,10;MQ=40;FQ=126;PV4=0.18,1,0.001,0.016 |
| 87972884 | A | G | 999 | DP=466;VDB=0.0421;AF1=1;AC1=6;DP4=0,0,283,145;MQ=49;FQ=-245 |
| 88049071 | C | T | 999 | DP=236;VDB=0.0044;AF1=1;AC1=6;DP4=0,0,55,126;MQ=38;FQ=-182 |
| 88051435 | A | G | 999 | DP=506;VDB=0.0459;AF1=1;AC1=6;DP4=0,1,161,324;MQ=38;FQ=-288;PV4=1,0.2,1,0.18 |
| 88083371 | CTTTT | CTT | 18.8 | INDEL;DP=55;VDB=0.0510;AF1=0.1667;AC1=1;DP4=7,41,1,1;MQ=49;FQ=19.5;PV4=0.3,1,1,0.096 |
| 88091016 | T | TG | 11.1 | INDEL;DP=22;VDB=0.0416;AF1=0.1675;AC1=1;DP4=5,13,0,2;MQ=49;FQ=11.8;PV4=1,1,0,0.0016 |
| 88101194 | T | A | 44.5 | DP=75;VDB=0.0000;AF1=0.5571;AC1=3;DP4=1,2,8,0;MQ=40;FQ=47.6;PV4=0.055,4.6e-06,0.00014,0.00055 |
| 88101195 | T | A | 35.9 | DP=69;VDB=0.0000;AF1=0.5481;AC1=3;DP4=1,2,6,0;MQ=42;FQ=39.3;PV4=0.083,0.00041,0,0.00096 |

| | | | | |
|---|---|---|---|---|
| 88130237 | G | GCC | 170 | INDEL;DP=88;VDB=0.0009;AF1=0.5;G3=1.555e-14,1,3.505e-22;HWE=0.0414;AC1=3;DP4=3,14,2,11;MQ=48;FQ=173;PV4=1,0.00053,0.34,0.00053 |
| 88132037 | T | TC | 36.9 | INDEL;DP=154;VDB=0.0092;AF1=0.1754;AC1=1;DP4=10,9,6,0;MQ=49;FQ=37.8;PV4=0.057,1.5e-05,0.11,0.0035 |
| 88134642 | G | T | 999 | DP=58;VDB=0.0457;AF1=1;AC1=6;DP4=1,0,12,27;MQ=50;FQ=-49.8;PV4=0.32,0.0017,1,1 |
| **88136484** | **G** | **C** | **999** | **DP=318;VDB=0.0494;AF1=1;AC1=6;DP4=0,0,177,124;MQ=40;FQ=-269** |
| 88136812 | A | AG | 28.6 | INDEL;DP=40;VDB=0.0257;AF1=0.2839;AC1=2;DP4=20,7,4,1;MQ=49;FQ=30;PV4=1,1,0.2,0.0099 |
| 88138840 | TCTCAC | TC | 156 | INDEL;DP=48;VDB=0.0509;AF1=0.5063;AC1=3;DP4=2,0,3,4;MQ=42;FQ=-6.26;PV4=0.44,0.28,1,1 |
| 88138983 | AGT | A | 7.54 | INDEL;DP=71;VDB=0.0453;AF1=0.1667;AC1=1;DP4=40,26,1,1;MQ=49;FQ=8.2;PV4=1,0.17,1,1 |
| 88146362 | G | A | 999 | DP=62;VDB=0.0001;AF1=1;AC1=6;DP4=0,0,10,13;MQ=38;FQ=-48.9 |
| 88154757 | G | A | 61.9 | DP=36;VDB=0.0000;AF1=1;AC1=6;DP4=0,0,2,4;MQ=37;FQ=-34 |
| 88154897 | T | A | 20.8 | DP=38;VDB=0.0000;AF1=1;AC1=6;DP4=0,0,3,1;MQ=46;FQ=-30.4 |
| 88159661 | A | AG | 47.8 | INDEL;DP=104;VDB=0.0077;AF1=0.3341;AC1=2;DP4=9,15,7,0;MQ=45;FQ=49.5;PV4=0.0068,0.0057,7.5e-05,0.00042 |
| 88199735 | G | GC | 10.5 | INDEL;DP=39;VDB=0.0438;AF1=0.3883;AC1=2;DP4=12,14,5,0;MQ=50;FQ=12.4;PV4=0.048,0.011,1,0.0012 |
| 88203629 | TC | TCC | 18.6 | INDEL;DP=81;VDB=0.0391;AF1=0.3768;AC1=2;DP4=14,14,0,6;MQ=45;FQ=20.5;PV4=0.031,0.11,0.0012,0.00043 |
| 88215576 | ATTTTTTTTTTTTTT | ATTTTTTTTTTTTTTT,ATTTTTTTTTTTTTTTT,ATTTTTTTTTTTTTTTT | 62.7 | INDEL;DP=271;VDB=0.0500;AF1=1;AC1=6;DP4=4,7,37,41;MQ=49;FQ=-71.6;PV4=0.54,1,0.17,0.042 |
| 88224499 | G | A | 999 | DP=350;VDB=0.0512;AF1=1;AC1=6;DP4=0,0,163,176;MQ=42;FQ=-288 |
| 88241176 | T | TA | 136 | INDEL;DP=27;VDB=0.0069;AF1=0.7962;AC1=5;DP4=1,1,12,1;MQ=39;FQ=8.27;PV4=0.26,0.0036,0.0021,0.0014 |
| 88241825 | GA | G | 28.7 | INDEL;DP=27;VDB=0.0090;AF1=0.4998;AC1=3;DP4=1,2,5,0;MQ=48;FQ=31.7;PV4=0.11,0.0038,0.24,0.0097 |
| 88243665 | CAGAGA | CAGA | 48.1 | INDEL;DP=26;VDB=0.0326;AF1=0.1849;AC1=1;DP4=8,10,0,3;MQ=50;FQ=49;PV4=0.26,0.052,1,1 |
| 88244799 | CA | CAA | 106 | INDEL;DP=66;VDB=0.0091;AF1=0.4989;G3=3.894e-10,1,6.157e-13;HWE=0.0419;AC1=3;DP4=8,8,7,2;MQ=45;FQ=109;PV4=0.23,0.34,0.00022,0.00027 |

| | | | | |
|---|---|---|---|---|
| 88253599 | C | CT | 13.7 | INDEL;DP=36;VDB=0.0008;AF1=0.399;AC1=2;DP4=4,0,0,3;MQ=42;FQ=15.9;PV4=0.029,0.016,0.13,0.12 |
| 88257055 | T | A | 21.2 | DP=41;VDB=0.0001;AF1=0.4841;G3=6.951e-08,1,4.209e-14;HWE=0.0475;AC1=3;DP4=2,3,0,6;MQ=39;FQ=24;PV4=0.18,1.7e-08,1,0.00018 |
| 88285855 | GAAAAAAAAAAAAA | GAAAAAAAAAAAAA,GAAAAAAAAAAA | 27.7 | INDEL;DP=494;VDB=0.0375;AF1=1;AC1=6;DP4=17,10,98,61;MQ=46;FQ=-66.6;PV4=1,1,0.38,0.27 |
| 88331385 | G | A | 999 | DP=21;VDB=0.0367;AF1=1;AC1=6;DP4=0,0,5,12;MQ=48;FQ=-44 |
| 88333638 | T | C | 999 | DP=20;VDB=0.0342;AF1=1;AC1=6;DP4=0,0,3,14;MQ=49;FQ=-44 |
| 88341855 | T | C | 999 | DP=265;VDB=0.0509;AF1=1;AC1=6;DP4=0,0,127,116;MQ=47;FQ=-275 |
| 88343716 | TAAAAAAAAAAA | TAAAAAAAAAAAAA,TAAAAAAAAAAAAAA,TAAAAAAAAAAAAA | 266 | INDEL;DP=52;VDB=0.0239;AF1=1;AC1=6;DP4=0,0,14,30;MQ=48;FQ=-53.4 |
| 88346621 | A | AGC | 9.27 | INDEL;DP=152;VDB=0.0034;AF1=0.2767;AC1=2;DP4=14,12,2,3;MQ=44;FQ=10.4;PV4=0.65,0.0084,0.016,0.0027 |
| 88346705 | T | A | 999 | DP=735;VDB=0.0001;AF1=1;AC1=6;DP4=2,5,441,150;MQ=32;FQ=-260;PV4=0.015,2.3e-05,1.5e-07,1 |
| 88346708 | C | A | 9.64 | DP=447;VDB=0.0000;AF1=0.4407;AC1=3;DP4=2,4,5,0;MQ=43;FQ=11.8;PV4=0.061,7.2e-06,0.00025,0.0013 |
| 88359256 | A | AG | 32.6 | INDEL;DP=22;VDB=0.0390;AF1=0.2033;AC1=1;DP4=7,3,1,2;MQ=43;FQ=33.5;PV4=0.51,0.2,0.079,0.005 |
| 88363699 | G | T | 17.2 | DP=86;VDB=0.0001;AF1=0.4831;G3=5.611e-08,1,3.522e-60;HWE=0.0468;AC1=3;DP4=7,9,9,1;MQ=45;FQ=20;PV4=0.037,2.6e-11,9.3e-05,0.00034 |
| 88367545 | T | TA | 999 | INDEL;DP=32;VDB=0.0440;AF1=1;AC1=6;DP4=0,0,15,15;MQ=50;FQ=-59.5 |
| 88367708 | C | A | 999 | DP=67;VDB=0.0458;AF1=1;AC1=6;DP4=0,0,32,31;MQ=44;FQ=-78.9 |
| 88369870 | G | A | 999 | DP=24;VDB=0.0495;AF1=1;AC1=6;DP4=0,0,8,8;MQ=47;FQ=-43.8 |
| 88383614 | T | G | 53.3 | DP=41;VDB=0.0026;AF1=0.5005;G3=1.155e-09,1,1.848e-08;HWE=0.0418;AC1=3;DP4=11,0,8,0;MQ=34;FQ=56.3;PV4=1,3.1e-07,0.22,0.0066 |
| 88383672 | A | G | 999 | DP=27;VDB=0.0482;AF1=1;AC1=6;DP4=0,0,22,2;MQ=33;FQ=-48.4 |
| 88383688 | TGGGC | TGGGCGGTTCTCGGGGC | 242 | INDEL;DP=29;VDB=0.0430;AF1=1;AC1=6;DP4=0,0,4,8;MQ=29;FQ=-44 |
| 88383703 | G | C | 249 | DP=29;VDB=0.0390;AF1=1;AC1=6;DP4=0,0,10,9;MQ=35;FQ=-36.9 |

| | | | | |
|---|---|---|---|---|
| 88387698 | G | C | 16.6 | DP=604;VDB=0.0001;AF1=0.4952;G3=6.962e-08,1,1.53e-46;HWE=0.0429;AC1=3;DP4=11,18,3,8;MQ=41;FQ=19.5;PV4=0.72,5.3e-19,3e-05,0.00039 |
| 88390865 | G | C | 204 | DP=321;VDB=0.0000;AF1=1;AC1=6;DP4=0,0,5,35;MQ=37;FQ=-54.9 |
| 88392125 | G | A | 999 | DP=3605;VDB=0.0504;AF1=1;AC1=6;DP4=0,4,1251,2211;MQ=48;FQ=-288;PV4=0.3,7.7e-07,1,1 |
| 88439645 | G | A | 87 | DP=20;VDB=0.0004;AF1=0.8289;AC1=5;DP4=0,1,10,0;MQ=34;FQ=-21.7;PV4=0.091,0.00024,1,1 |
| 88443241 | A | T | 42.9 | DP=52;VDB=0.0001;AF1=0.6536;AC1=4;DP4=4,1,1,7;MQ=37;FQ=47.5;PV4=0.032,4.4e-08,4.6e-20,0.09 |
| 88443243 | A | T | 201 | DP=57;VDB=0.0002;AF1=0.6666;AC1=4;DP4=4,0,7,17;MQ=33;FQ=48;PV4=0.016,6.1e-12,2.3e-06,0.00061 |
| 88506469 | A | T | 6.43 | DP=25;VDB=0.0145;AF1=0.1693;AC1=1;DP4=12,11,1,1;MQ=50;FQ=7.07;PV4=1,1,6.5e-13,1 |
| 88507229 | A | G | 16.1 | DP=21;VDB=0.0108;AF1=0.3493;AC1=2;DP4=13,2,4,1;MQ=48;FQ=18;PV4=1,2.2e-05,0.18,0.49 |
| 88512513 | GTAT | GT | 15.8 | INDEL;DP=37;VDB=0.0473;AF1=0.1669;AC1=1;DP4=27,8,1,1;MQ=50;FQ=16.6;PV4=0.43,0.12,1,1 |
| 88533031 | A | C | 11.2 | DP=73;VDB=0.0000;AF1=0.4687;AC1=3;DP4=3,5,3,4;MQ=43;FQ=13.7;PV4=1,1.6e-08,1,0.0093 |
| 88667535 | G | A | 999 | DP=25;VDB=0.0002;AF1=0.981;AC1=5;DP4=0,1,13,9;MQ=37;FQ=-28.6;PV4=0.43,0.0052,1,1 |
| 88691567 | TG | T | 45 | INDEL;DP=27;VDB=0.0002;AF1=0.3599;AC1=2;DP4=0,3,2,2;MQ=29;FQ=12.4;PV4=0.43,0.31,1,0.0033 |
| 88703850 | AGG | AGGG | 6.46 | INDEL;DP=37;VDB=0.0194;AF1=0.1942;AC1=1;DP4=3,4,1,1;MQ=34;FQ=7.22;PV4=1,1,0.39,0.018 |
| 88706673 | C | CG | 81.2 | INDEL;DP=29;VDB=0.0184;AF1=0.3537;AC1=2;DP4=8,3,2,6;MQ=41;FQ=83.2;PV4=0.07,0.46,0.085,0.00016 |
| 88712975 | ATGTGTGTGTGTGTGTGTGTGTGTGTGTG | ATGTGTGTGTGTGTGTGTGTGTGTGTGTG | 25.2 | INDEL;DP=34;VDB=0.0456;AF1=1;AC1=6;DP4=0,0,0,2;MQ=50;FQ=-35 |
| 88714342 | TG | TGG | 170 | INDEL;DP=88;VDB=0.0002;AF1=1;AC1=6;DP4=0,0,31,0;MQ=43;FQ=-62 |
| 88716701 | C | CA | 36.4 | INDEL;DP=30;VDB=0.0119;AF1=0.4926;AC1=3;DP4=1,1,4,1;MQ=37;FQ=7.2;PV4=1,0.013,0.27,0.0012 |
| 88716851 | TG | T | 120 | INDEL;DP=42;VDB=0.0001;AF1=0.3459;AC1=2;DP4=9,1,9,1;MQ=36;FQ=122;PV4=1,1,0.24,1 |

| | | | | |
|---|---|---|---|---|
| 88737484 | T | TC | 38.6 | INDEL;DP=56;VDB=0.0027;AF1=0.2049;AC1=1;DP4=5,1,0,9;MQ=40;FQ=39.6;PV4=0.002,0.12,4.6e-06,0.00095 |
| 88740178 | A | T | 10.3 | DP=51;VDB=0.0000;AF1=0.488;G3=1.304e-07,1,2.007e-18;HWE=0.0451;AC1=3;DP4=4,4,2,4;MQ=42;FQ=13;PV4=0.63,7.7e-10,0.00062,0.0022 |
| 88805795 | AGG | AG | 999 | INDEL;DP=23;VDB=0.0017;AF1=1;AC1=6;DP4=0,0,12,3;MQ=35;FQ=-43.3 |
| 88855908 | TA | T | 48.8 | INDEL;DP=25;VDB=0.0005;AF1=0.6572;AC1=4;DP4=0,1,4,0;MQ=48;FQ=-23;PV4=0.2,1,0.34,0.0011 |
| 88862563 | G | A | 6.62 | DP=38;VDB=0.0000;AF1=0.5098;AC1=3;DP4=1,3,3,1;MQ=41;FQ=9.22;PV4=0.49,0.013,0.067,0.012 |
| 88867840 | G | GAA | 152 | INDEL;DP=36;VDB=0.0001;AF1=1;AC1=6;DP4=0,0,8,0;MQ=36;FQ=-39.5 |
| 88870344 | TGAG | TG | 97.8 | INDEL;DP=101;VDB=0.0004;AF1=0.1718;AC1=1;DP4=5,11,6,3;MQ=44;FQ=98.7;PV4=0.12,0.14,0.0048,1 |
| 88871074 | TA | T | 170 | INDEL;DP=21;VDB=0.0169;AF1=0.4992;G3=2.451e-09,1,3.794e-10;HWE=0.0424;AC1=3;DP4=4,1,9,2;MQ=39;FQ=86.9;PV4=1,0.055,0.00023,0.00022 |
| 88874570 | T | G | 26.9 | DP=26;VDB=0.0009;AF1=1;AC1=6;DP4=0,0,1,3;MQ=42;FQ=-30.3 |
| 88904754 | CT | C | 72.3 | INDEL;DP=72;VDB=0.0039;AF1=0.3421;AC1=2;DP4=9,25,8,3;MQ=28;FQ=74.1;PV4=0.011,0.19,1,0.1 |
| 88973538 | AG | A | 72.7 | INDEL;DP=83;VDB=0.0480;AF1=0.4997;G3=1.219e-09,1,3.741e-71;HWE=0.0415;AC1=3;DP4=26,26,4,7;MQ=42;FQ=75.7;PV4=0.51,1,0.012,0.00013 |
| 88979953 | TAAAA | TAA | 36.7 | INDEL;DP=92;VDB=0.0407;AF1=0.1667;AC1=1;DP4=30,55,2,2;MQ=49;FQ=37.5;PV4=0.62,1,1,0.024 |
| 88984811 | GTA | G | 999 | INDEL;DP=143;VDB=0.0003;AF1=0.5129;G3=2.009e-57,1,2.131e-07;HWE=0.0455;AC1=3;DP4=2,8,43,20;MQ=43;FQ=75.2;PV4=0.0055,0.11,0.0016,0.35 |
| 88992067 | G | GGC | 44.5 | INDEL;DP=21;VDB=0.0348;AF1=0.4001;AC1=2;DP4=6,5,3,0;MQ=47;FQ=46.7;PV4=0.26,0.45,1,0.0017 |
| 89009367 | A | AAC | 111 | INDEL;DP=72;VDB=0.0034;AF1=0.5998;AC1=4;DP4=2,1,0,7;MQ=39;FQ=80.7;PV4=0.067,0.07,0.00078,0.034 |
| 89012351 | G | A | 90.4 | DP=87;VDB=0.0001;AF1=0.5771;AC1=3;DP4=1,3,3,16;MQ=38;FQ=29.2;PV4=1,7.2e-15,0.13,0.00067 |
| 89020808 | AGG | AG | 142 | INDEL;DP=68;VDB=0.0111;AF1=0.484;G3=6.448e-07,1,2.651e-25;HWE=0.0465;AC1=3;DP4=2,7,11,3;MQ=40;FQ=142;PV4=0.013,0.05,0.17,0.00077 |

| | | | | |
|---|---|---|---|---|
| 89025032 | G | GT | 999 | INDEL;DP=80;VDB=0.0007;AF1=1;AC1=6;DP4=0,1,28,21;MQ=36;FQ=-61.4;PV4=0.44,0.4,0.11,0.0036 |
| 89052995 | G | GA | 202 | INDEL;DP=26;VDB=0.0211;AF1=0.6285;AC1=4;DP4=9,3,0,11;MQ=41;FQ=207;PV4=0.00034,1,0.1,0.27 |
| 89072521 | C | CG | 162 | INDEL;DP=70;VDB=0.0007;AF1=0.501;G3=4.902e-15,1,3.209e-08;HWE=0.0417;AC1=3;DP4=2,2,5,3;MQ=44;FQ=53.8;PV4=1,0.28,0.049,0.0049 |
| 89148133 | C | A | 6.46 | DP=111;VDB=0.0001;AF1=0.1709;AC1=1;DP4=6,11,1,2;MQ=45;FQ=7.11;PV4=1,0.00056,0.012,0.046 |
| 89149862 | C | CG | 4.57 | INDEL;DP=53;VDB=0.0085;AF1=0.1672;AC1=1;DP4=15,7,2,0;MQ=42;FQ=5.11;PV4=1,1,0.26,0.1 |
| 89228662 | CA | C | 8.39 | INDEL;DP=22;VDB=0.0332;AF1=0.1675;AC1=1;DP4=8,11,2,0;MQ=49;FQ=9.08;PV4=0.21,0.49,1,1 |
| 89236368 | C | CT | 5.86 | INDEL;DP=29;VDB=0.0376;AF1=0.3557;AC1=2;DP4=1,9,0,3;MQ=47;FQ=7.3;PV4=1,0.4,1.4e-06,0.00071 |
| 89249603 | T | TA | 3.16 | INDEL;DP=52;VDB=0.0188;AF1=0.2674;AC1=2;DP4=6,8,3,0;MQ=48;FQ=3.86;PV4=0.21,0.02,0,0.0094 |
| 89250197 | A | AG | 16.6 | INDEL;DP=22;VDB=0.0194;AF1=0.47;AC1=3;DP4=5,3,0,5;MQ=44;FQ=19.3;PV4=0.075,0.13,0.024,0.00012 |
| 89250334 | G | GC | 16.7 | INDEL;DP=22;VDB=0.0301;AF1=0.2606;AC1=1;DP4=3,3,2,1;MQ=45;FQ=18;PV4=1,0.0049,2.2e-05,0.048 |
| 89252189 | T | TC | 191 | INDEL;DP=63;VDB=0.0003;AF1=0.8138;AC1=5;DP4=1,1,10,0;MQ=43;FQ=10.5;PV4=0.17,1,0.087,0.02 |
| 89268369 | GG | GGCTG | 999 | INDEL;DP=169;VDB=0.0001;AF1=0.815;AC1=5;DP4=2,2,46,0;MQ=30;FQ=-19.9;PV4=0.0049,1,5.5e-10,0.0012 |
| 89544201 | GT | GTTT | 46.5 | INDEL;DP=44;VDB=0.0022;AF1=0.6283;AC1=4;DP4=0,2,10,0;MQ=39;FQ=26.7;PV4=0.015,1.2e-07,0,0.00012 |
| 89551090 | G | A | 120 | DP=35;VDB=0.0001;AF1=0.6664;AC1=4;DP4=2,3,1,12;MQ=39;FQ=92.5;PV4=0.17,7.5e-08,0.00057,0.00088 |
| 89554154 | C | CT | 46.9 | INDEL;DP=40;VDB=0.0262;AF1=0.4748;G3=7.256e-07,1,4.219e-34;HWE=0.0498;AC1=3;DP4=7,2,0,8;MQ=42;FQ=49.7;PV4=0.0023,0.00041,0.15,0.0036 |
| 89564934 | G | A | 3.71 | DP=25;VDB=0.0132;AF1=0.1689;AC1=1;DP4=11,12,0,2;MQ=49;FQ=4.19;PV4=0.49,1,0.06,1 |
| 89583491 | TG | TGG | 123 | INDEL;DP=51;VDB=0.0001;AF1=1;AC1=6;DP4=0,0,9,1;MQ=29;FQ=-45.5 |
| 89758216 | G | C | 999 | DP=60;VDB=0.0002;AF1=0.8329;AC1=5;DP4=1,2,13,19;MQ=39;FQ=26.1;PV4=1,3.7e-07,0.024,0.03 |
| 89758521 | C | CA | 37 | INDEL;DP=24;VDB=0.0115;AF1=0.7018;AC1=4;DP4=2,1,3,1;MQ=45;FQ=41.6;PV4=1,0.033,0.026,0.038 |

| | | | | |
|---|---|---|---|---|
| 89819222 | A | T | 11.3 | DP=27;VDB=0.0000;AF1=0.6148;AC1=4;DP4=2,3,7,0;MQ=36;FQ=15;PV4=0.045,0.00053,0.008,0.00025 |
| 89820618 | TA | T | 149 | INDEL;DP=102;VDB=0.0001;AF1=0.5999;AC1=4;DP4=4,1,3,6;MQ=41;FQ=67.7;PV4=0.27,0.08,0.079,0.00014 |
| 89823268 | G | GCC | 999 | INDEL;DP=60;VDB=0.0002;AF1=0.6667;AC1=4;DP4=2,2,20,0;MQ=39;FQ=46.8;PV4=0.022,0.00024,0.00071,0.0003 |
| 89847735 | G | GT | 999 | INDEL;DP=98;VDB=0.0000;AF1=1;AC1=6;DP4=1,0,14,0;MQ=35;FQ=-37;PV4=1,1,0.32,1 |
| 89848775 | C | A | 999 | DP=166;VDB=0.0001;AF1=1;AC1=6;DP4=0,1,67,58;MQ=34;FQ=-116;PV4=0.47,7.2e-08,0.0017,0.00018 |
| 89862405 | A | T | 180 | DP=53;VDB=0.0000;AF1=0.5;G3=1.789e-13,1,1.467e-14;HWE=0.0414;AC1=3;DP4=7,3,12,0;MQ=41;FQ=183;PV4=0.078,0.028,0.03,0.0016 |
| 89869429 | C | T | 10.9 | DP=34;VDB=0.0001;AF1=0.1785;AC1=1;DP4=8,13,0,4;MQ=47;FQ=11.7;PV4=0.27,4.9e-06,1.6e-05,0.0049 |
| 89903449 | G | T | 26.5 | DP=159;VDB=0.0002;AF1=0.3131;AC1=2;DP4=119,16,7,9;MQ=29;FQ=28;PV4=0.00012,1,1,1 |
| 90058481 | G | T | 191 | DP=65;VDB=0.0002;AF1=0.5;G3=1.472e-13,1,1.854e-23;HWE=0.0414;AC1=3;DP4=7,9,9,13;MQ=39;FQ=194;PV4=1,1.7e-08,0.00038,1 |
| 90062192 | TTGTGTGTGTGTGTGTGTGTCTGTGTGTGTGTGTGTGT | TTGTGTGTGTGTGTGTGTCTGTGTGTGTGTGTGT | 84.4 | INDEL;DP=67;VDB=0.0486;AF1=0.4872;G3=2.047e-07,1,1.064e-126;HWE=0.0454;AC1=3;DP4=19,17,3,4;MQ=48;FQ=87.2;PV4=0.7,0.12,0.34,0.44 |
| 90155494 | CCA | CCACA | 999 | INDEL;DP=110;VDB=0.0002;AF1=0.898;AC1=5;DP4=3,2,52,0;MQ=44;FQ=-33.2;PV4=0.0063,1,0.27,0.0081 |
| 90191372 | C | CG | 47 | INDEL;DP=52;VDB=0.0005;AF1=0.8116;AC1=5;DP4=1,0,8,0;MQ=43;FQ=-14;PV4=1,8.2e-06,0.21,0.022 |
| 90196813 | GCT | GCTCT | 43.8 | INDEL;DP=63;VDB=0.0216;AF1=0.4979;G3=2.659e-09,1,6.069e-29;HWE=0.042;AC1=3;DP4=8,10,7,0;MQ=45;FQ=46.8;PV4=0.02,2.7e-08,0.036,0.0018 |
| 90199059 | C | G | 21.2 | DP=20;VDB=0.0009;AF1=0.356;AC1=2;DP4=2,2,1,4;MQ=36;FQ=23;PV4=0.52,1.3e-05,6.2e-05,0.0084 |
| 90206228 | A | C | 999 | DP=41;VDB=0.0001;AF1=0.8535;AC1=5;DP4=2,0,23,6;MQ=37;FQ=-21;PV4=1,2.6e-05,0.0016,0.00012 |
| 90207904 | T | G | 29.4 | DP=45;VDB=0.0001;AF1=0.7721;AC1=5;DP4=0,1,0,4;MQ=32;FQ=4.6;PV4=1,0.013,0,0.00034 |
| 90207906 | C | T | 127 | DP=61;VDB=0.0001;AF1=0.8246;AC1=5;DP4=1,1,2,12;MQ=35;FQ=24.3;PV4=0.35,2.6e-05,0.075,0.044 |
| 90208527 | G | C | 88.7 | DP=70;VDB=0.0001;AF1=1;AC1=6;DP4 |

| | | | | |
|---|---|---|---|---|
| | | | | =0,0,0,11;MQ=38;FQ=-39.3 |
| 90217450 | A | C | 20.3 | DP=32;VDB=0.0001;AF1=1;AC1=6;DP4=0,0,1,2;MQ=30;FQ=-29.8 |
| 90219075 | CTT | CT | 34.2 | INDEL;DP=24;VDB=0.0011;AF1=0.5406;AC1=3;DP4=1,1,0,3;MQ=42;FQ=15.6;PV4=0.4,0.4,0.00026,0.0029 |
| 90219529 | CT | CTT | 8.02 | INDEL;DP=39;VDB=0.0022;AF1=0.3032;AC1=2;DP4=1,2,2,2;MQ=37;FQ=9.45;PV4=1,0.18,0.071,0.035 |
| 90225742 | C | A | 62.5 | DP=38;VDB=0.0001;AF1=0.4949;G3=9.858e-08,1,1.535e-14;HWE=0.043;AC1=3;DP4=0,10,9,1;MQ=44;FQ=65.5;PV4=0.00012,4e-09,0,0.0016 |
| 90227921 | C | G | 42.3 | DP=26;VDB=0.0076;AF1=0.5773;AC1=4;DP4=1,0,1,1;MQ=46;FQ=8.44;PV4=1,0.44,0,1 |
| 90259941 | GC | G | 999 | INDEL;DP=39;VDB=0.0032;AF1=0.5005;G3=1.676e-34,1,4.875e-09;HWE=0.0416;AC1=3;DP4=1,4,9,11;MQ=38;FQ=66.7;PV4=0.61,0.24,0.013,0.00018 |
| 90261001 | TG | T | 78.6 | INDEL;DP=51;VDB=0.0435;AF1=0.3256;AC1=2;DP4=8,10,4,3;MQ=47;FQ=80.3;PV4=0.67,4e-05,0.00046,0.0014 |
| 90284521 | TC | T | 65.7 | INDEL;DP=21;VDB=0.0259;AF1=0.4997;G3=1.15e-09,1,9.448e-12;HWE=0.0415;AC1=3;DP4=3,2,0,7;MQ=39;FQ=68.7;PV4=0.045,1,0.0046,0.002 |
| 90289608 | TCAG | T | 31.3 | INDEL;DP=157;VDB=0.0080;AF1=0.1934;AC1=1;DP4=20,21,7,0;MQ=45;FQ=32.3;PV4=0.014,1.2e-10,0.0073,0.00012 |
| 90293519 | T | TG | 4.91 | INDEL;DP=44;VDB=0.0001;AF1=1;AC1=6;DP4=0,0,2,0;MQ=37;FQ=-35 |
| 90307207 | G | T | 67.7 | DP=126;VDB=0.0001;AF1=0.3893;AC1=2;DP4=34,20,14,3;MQ=37;FQ=69.9;PV4=0.23,0.028,1,0.026 |
| 90319968 | G | GT | 15.3 | INDEL;DP=156;VDB=0.0001;AF1=0.6364;AC1=4;DP4=1,1,6,0;MQ=36;FQ=14.9;PV4=0.25,1.8e-08,0.002,0.012 |
| 90328271 | T | TC | 270 | INDEL;DP=33;VDB=0.0001;AF1=1;AC1=6;DP4=0,0,4,18;MQ=34;FQ=-49.5 |
| 90330823 | TAA | TA | 13.9 | INDEL;DP=55;VDB=0.0459;AF1=0.1675;AC1=1;DP4=15,3,1,1;MQ=49;FQ=14.7;PV4=0.37,1,1,0.04 |
| 90343949 | TGAGAGAGAGAGAGAGAGAGAGAGAGA | TGAGAGAGAGAGAGAGAGAGAGA | 92 | INDEL;DP=30;VDB=0.0363;AF1=0.3432;AC1=2;DP4=8,7,1,3;MQ=50;FQ=93.8;PV4=0.58,1,1,1 |
| 90404196 | T | G | 19.4 | DP=41;VDB=0.0003;AF1=0.3342;AC1=2;DP4=27,2,8,0;MQ=42;FQ=21.1;PV4=1,8.7e-07,0.045,0.00069 |
| 90455797 | A | G | 999 | DP=668;VDB=0.0000;AF1=1;AC1=6;DP4=1,1,1,639;MQ=32;FQ=-222;PV4=0.0062,2.4e-05,0.026,0.011 |
| 90455816 | A | G | 999 | DP=668;VDB=0.0000;AF1=1;AC1=6;DP4=1,2,1,654;MQ=32;FQ=-288;PV4=0.0091,1,0.062,1 |

| | | | | |
|---|---|---|---|---|
| 90692497 | TA | T | 188 | INDEL;DP=85;VDB=0.0006;AF1=0.4935;G3=2.454e-07,1,6.219e-21;HWE=0.0434;AC1=3;DP4=1,9,10,3;MQ=46;FQ=191;PV4=0.0028,4.4e-08,0.0074,0.00013 |
| 92190533 | A | T | 69.1 | DP=59;VDB=0.0000;AF1=0.8016;AC1=5;DP4=1,0,12,0;MQ=48;FQ=-11.3;PV4=1,1.5e-07,0.35,1 |
| 92241511 | A | G | 65.4 | DP=23;VDB=0.0042;AF1=1;AC1=6;DP4=0,0,0,9;MQ=50;FQ=-35.5 |
| 92415221 | TCC | TC | 35 | INDEL;DP=21;VDB=0.0310;AF1=0.3426;AC1=2;DP4=12,4,3,1;MQ=49;FQ=36.9;PV4=1,1,0.27,0.00073 |
| 92494860 | TC | T | 10.4 | INDEL;DP=31;VDB=0.0512;AF1=0.6976;AC1=4;DP4=4,0,3,0;MQ=30;FQ=-11.4;PV4=1,1,1,0.14 |
| 92554944 | C | CAA | 33.8 | INDEL;DP=24;VDB=0.0282;AF1=0.3942;AC1=2;DP4=4,7,1,3;MQ=47;FQ=36;PV4=1,1.6e-13,1,0.021 |
| 92822563 | T | TCA | 99.7 | INDEL;DP=369;VDB=0.0001;AF1=0.4999;G3=2.33e-11,1,6.639e-63;HWE=0.0414;AC1=3;DP4=2,71,2,15;MQ=41;FQ=103;PV4=0.16,0.00022,1,0.00085 |
| 93038935 | A | G | 101 | DP=20;VDB=0.0381;AF1=1;AC1=6;DP4=0,0,15,0;MQ=14;FQ=-41.9 |
| 93038939 | T | C | 136 | DP=21;VDB=0.0406;AF1=1;AC1=6;DP4=0,0,20,0;MQ=17;FQ=-44.9 |
| 93039686 | A | T | 177 | DP=57;VDB=0.0000;AF1=1;AC1=6;DP4=0,0,3,10;MQ=30;FQ=-38.8 |
| 93040051 | A | G | 211 | DP=23;VDB=0.0435;AF1=1;AC1=6;DP4=0,0,17,5;MQ=19;FQ=-50.2 |
| 93040056 | A | G | 167 | DP=26;VDB=0.0497;AF1=0.85;AC1=5;DP4=1,0,19,5;MQ=21;FQ=-20;PV4=1,8.3e-06,0.00012,1 |
| 93040058 | T | C | 161 | DP=26;VDB=0.0497;AF1=0.8498;AC1=5;DP4=1,0,18,5;MQ=21;FQ=-20;PV4=1,1.7e-06,0.00017,1 |
| 93040065 | T | C | 140 | DP=25;VDB=0.0453;AF1=0.8379;AC1=5;DP4=1,0,16,5;MQ=22;FQ=-15.1;PV4=1,2.3e-07,0.00038,1 |
| 93040077 | T | C | 155 | DP=21;VDB=0.0346;AF1=0.8353;AC1=5;DP4=1,0,14,3;MQ=21;FQ=-14.1;PV4=1,0.26,0.00027,0.28 |
| 93040122 | A | G | 30.1 | DP=49;VDB=0.0030;AF1=0.7105;AC1=5;DP4=1,0,7,0;MQ=28;FQ=-11.1;PV4=1,0.00064,0.019,1 |
| 93040125 | T | C | 29.2 | DP=48;VDB=0.0006;AF1=0.7216;AC1=5;DP4=1,0,9,0;MQ=27;FQ=-16.7;PV4=1,2.7e-06,0.0082,1 |
| 93040310 | T | A | 75.6 | DP=83;VDB=0.0044;AF1=1;AC1=6;DP4=0,0,0,7;MQ=28;FQ=-34.8 |
| 93082568 | T | C | 37.2 | DP=31;VDB=0.0005;AF1=0.5476;AC1=4;DP4=3,4,9,9;MQ=24;FQ=38.4;PV4=1,4.9e-05,0.0073,0.0076 |
| 93246389 | GGTA | G | 999 | INDEL;DP=687;VDB=0.0004;AF1=0.8417;AC1=5;DP4=6,11,177,73;MQ=29;FQ=-24.5;PV4=0.005,1.3e-14,1.4e-32,0.0008 |

| | | | | |
|---|---|---|---|---|
| 93253331 | AGCC | A | 27.1 | INDEL;DP=22;VDB=0.0162;AF1=0.4933;G3=9.607e-09,1,6.168e-15;HWE=0.0437;AC1=3;DP4=4,9,0,5;MQ=36;FQ=30;PV4=0.28,0.0025,0.29,1 |
| 93329166 | ATGTGTGTGTGTGTGT | ATGTGTGTGTGTGT | 12.2 | INDEL;DP=45;VDB=0.0421;AF1=0.3301;AC1=2;DP4=12,14,1,2;MQ=47;FQ=13.8;PV4=1,0.28,0.18,0.16 |
| 94148325 | C | G | 26.6 | DP=34;VDB=0.0001;AF1=0.6192;AC1=4;DP4=3,1,7,1;MQ=46;FQ=30.7;PV4=1,3.9e-05,0.3,0.029 |
| 94195130 | T | A | 213 | DP=32;VDB=0.0002;AF1=0.695;AC1=4;DP4=0,3,13,3;MQ=42;FQ=27.1;PV4=0.021,0.0033,0.065,0.33 |
| 94324191 | GAA | G | 7.58 | INDEL;DP=23;VDB=0.0060;AF1=0.4994;AC1=3;DP4=1,1,2,0;MQ=40;FQ=10.2;PV4=1,0.2,0.2,0.22 |
| 94418645 | C | CA | 999 | INDEL;DP=52;VDB=0.0035;AF1=0.8334;AC1=5;DP4=1,1,16,21;MQ=36;FQ=-6.59;PV4=1,0.41,9.6e-08,0.0011 |
| 94418749 | G | C | 8.1 | DP=51;VDB=0.0000;AF1=0.3317;AC1=2;DP4=13,5,1,6;MQ=44;FQ=9.61;PV4=0.021,1.2e-13,0.0064,0.00043 |
| 94432116 | T | G | 3.21 | DP=57;VDB=0.0001;AF1=0.3934;AC1=2;DP4=2,7,0,5;MQ=43;FQ=4.45;PV4=0.51,3.6e-07,9.2e-09,0.0013 |
| 94432415 | CA | CAGA | 15 | INDEL;DP=20;VDB=0.0434;AF1=0.3351;AC1=2;DP4=9,6,2,0;MQ=48;FQ=16.7;PV4=0.51,1,1,1 |
| 94454523 | T | A | 8.04 | DP=20;VDB=0.0132;AF1=0.1688;AC1=1;DP4=6,10,0,2;MQ=47;FQ=8.73;PV4=0.53,0.016,4.7e-05,1 |
| 94552257 | T | A | 20.3 | DP=20;VDB=0.0197;AF1=0.1683;AC1=1;DP4=7,10,1,1;MQ=48;FQ=21.1;PV4=1,1,0.18,0.35 |
| 94574625 | CT | CTT | 5.77 | INDEL;DP=24;VDB=0.0232;AF1=1;AC1=6;DP4=0,0,3,0;MQ=20;FQ=-35.7 |
| 94578123 | G | GGT | 999 | INDEL;DP=29;VDB=0.0206;AF1=0.6662;AC1=4;DP4=3,3,1,22;MQ=34;FQ=111;PV4=0.02,1,4.6e-08,0.00059 |
| 94676350 | G | C | 6.63 | DP=24;VDB=0.0000;AF1=0.7471;AC1=5;DP4=0,1,2,1;MQ=32;FQ=-9.2;PV4=1,0.014,1,0.025 |
| 94676351 | T | C | 5.89 | DP=24;VDB=0.0000;AF1=0.7467;AC1=5;DP4=0,1,2,1;MQ=32;FQ=-9.2;PV4=1,0.012,1,0.013 |
| 94747517 | GGTGTGTGTGTGTGTGTGTGTGTGTGTGT | GGTGTGTGTGTGTGTGTGTGTGT | 48.1 | INDEL;DP=48;VDB=0.0416;AF1=0.3651;AC1=2;DP4=14,5,3,3;MQ=46;FQ=50;PV4=0.34,1,1,1 |
| 94748323 | TC | T | 999 | INDEL;DP=188;VDB=0.0003;AF1=0.3333;AC1=2;DP4=12,64,5,34;MQ=40;FQ=999;PV4=0.79,1,0.014,1 |
| 94758173 | G | GCT | 69.8 | INDEL;DP=36;VDB=0.0364;AF1=0.4542;AC1=3;DP4=12,8,9,0;MQ=44;FQ=72.4;PV4=0.033,1,1.4e-10,0.0011 |
| 94760529 | G | A | 44.8 | DP=29;VDB=0.0001;AF1=0.3531;AC1=2;DP4=6,10,3,4;MQ=38;FQ=46.7;PV4=1,2.1e-10,0.089,0.014 |

| | | | | |
|---|---|---|---|---|
| 94761702 | A | T | 114 | DP=22;VDB=0.0087;AF1=0.4967;G3=1.551e-08,1,2.6e-23;HWE=0.0424;AC1=3;DP4=6,2,3,4;MQ=42;FQ=117;PV4=0.31,0.0066,0.0014,0.0025 |
| 94787901 | C | CG | 68.2 | INDEL;DP=23;VDB=0.0251;AF1=0.1853;AC1=1;DP4=7,3,0,7;MQ=43;FQ=69.1;PV4=0.0098,0.076,0.00018,0.0011 |
| 94811990 | T | TGG | 999 | INDEL;DP=22;VDB=0.0148;AF1=0.6629;AC1=4;DP4=2,1,0,14;MQ=44;FQ=22.2;PV4=0.022,0.4,0.43,0.002 |
| 94812840 | CT | C | 43.7 | INDEL;DP=57;VDB=0.0014;AF1=0.3935;AC1=3;DP4=2,10,2,2;MQ=46;FQ=45.7;PV4=0.24,0.42,1,0.23 |
| 94814263 | T | A | 141 | DP=86;VDB=0.0000;AF1=0.8861;AC1=5;DP4=0,1,2,17;MQ=43;FQ=-24.9;PV4=1,2.3e-07,0.22,0.00086 |
| 94868252 | C | T | 11.8 | DP=40;VDB=0.0000;AF1=0.3337;AC1=2;DP4=10,2,1,5;MQ=45;FQ=13.4;PV4=0.013,1e-07,0.00014,0.023 |
| 94871981 | A | AT | 18.2 | INDEL;DP=20;VDB=0.0159;AF1=0.5922;AC1=4;DP4=2,0,0,5;MQ=37;FQ=6.91;PV4=0.048,0.047,0.25,0.0043 |
| 94874934 | C | CA | 999 | INDEL;DP=40;VDB=0.0002;AF1=1;AC1=6;DP4=1,0,19,9;MQ=41;FQ=-40.6;PV4=1,0.18,0.087,0.011 |
| 94932334 | G | GC | 63.9 | INDEL;DP=21;VDB=0.0073;AF1=0.5966;AC1=4;DP4=1,2,4,1;MQ=46;FQ=40.8;PV4=0.46,0.27,0.13,0.089 |
| 94948565 | TG | TGTCAGG | 999 | INDEL;DP=736;VDB=0.0531;AF1=1;AC1=6;DP4=2,2,90,547;MQ=24;FQ=-296;PV4=0.1,1,5.4e-05,0.34 |
| 94968439 | A | C | 6.72 | DP=21;VDB=0.0076;AF1=0.2932;AC1=2;DP4=2,1,3,0;MQ=41;FQ=7.98;PV4=1,0.012,0.22,0.0029 |
| 94968440 | G | C | 5.23 | DP=20;VDB=0.0023;AF1=0.2931;AC1=2;DP4=2,1,3,0;MQ=41;FQ=6.37;PV4=1,0.00097,0.22,0.004 |
| 94993073 | TA | T | 71.9 | INDEL;DP=33;VDB=0.0455;AF1=0.3339;AC1=2;DP4=9,10,5,2;MQ=44;FQ=73.7;PV4=0.39,0.011,0.29,0.00043 |

# APPENDIX V

## Differential gene expression between *Matt*$^{ma/ma}$ and C57BL/6J using DSeq (v1.61)

## APPENDIX V

Complete list of 114 genes with different expression between F6 *Matt*$^{ma/ma}$ and C57BL6/J mice using DSeq (v1.61), with $p < 0.01$ over the entire genome. The only gene that mapped to the minimum region, chr3: 87,336,555 - 88,392,125, was *Transmembrane79* (highlighted in red). All genes were classified according to their Gene Ontology-associated biological processes.

| GeneName | Description | Chromosome | baseMean C57BL6/J | baseMean *Matt*$^{ma/ma}$ | log2 Fold Change | Gene Ontology |
|---|---|---|---|---|---|---|
| H2-Ea-ps | histocompatibility 2, class II antigen E alpha, pseudogene | 17 | 0.282 | 765 | 11.4 | Immune response |
| Gm18733 | predicted gene, 18733 | 17 | 8.91 | 207 | 4.54 | pseudogene |
| H2-Bl | histocompatibility 2, blastocyst | 17 | 14.7 | 213 | 3.86 | Immune response |
| Tpsab1 | tryptase alpha/beta 1 | 17 | 17.7 | 187 | 3.4 | Protein metabolic process |
| Npw | neuropeptide W | 17 | 4.41 | 42.8 | 3.28 | Signal transduction |
| RP23-381B19.7.1 | none | 11 | 79.8 | 620 | 2.96 | Unknown |
| Expi | extracellular proteinase inhibitor | 11 | 119 | 745 | 2.65 | Metabolic process |
| Gm10499 | predicted gene 10499 | 17 | 33.9 | 206 | 2.6 | pseudogene |
| Gm15921 | predicted gene 15921 | 5 | 26.2 | 135 | 2.36 | pseudogene |
| H2-K2 | histocompatibility 2, K region locus 2 | 17 | 135 | 649 | 2.26 | Immune response |
| Cd300lf | CD300 antigen like family member F | 11 | 11.8 | 55.5 | 2.23 | Cell differentation |
| C920025E04Rik | RIKEN cDNA C920025E04 gene | 17 | 122 | 545 | 2.16 | pseudogene |
| Wdfy1 | WD repeat and FYVE domain containing 1 | 1 | 690 | 2920 | 2.08 | Lipid metabolic process |
| 1500011B03Rik | RIKEN cDNA 1500011B03 gene | 5 | 47.1 | 193 | 2.03 | pseudogene |
| H2-Q1 | histocompatibility 2, Q region locus 1 | 17 | 429 | 1690 | 1.98 | Immune response |
| H2-Q10 | histocompatibility 2, Q region locus 10 | 17 | 197 | 668 | 1.76 | Immune response |
| 4933404O12Rik | RIKEN cDNA 4933404O12 gene | 5 | 261 | 870 | 1.74 | lincRNA |
| 2300002M23Rik | RIKEN cDNA 2300002M23 gene | 17 | 32.5 | 106 | 1.7 | Extracellular structure organization |
| Gm9000 | predicted gene 9000 | 13 | 9310 | 29500 | 1.66 | Unknown |
| Slc13a3 | solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 3 | 2 | 171 | 526 | 1.62 | Transmembrane transport |
| Il12rb1 | interleukin 12 receptor, | 8 | 44.1 | 133 | 1.6 | Immune |

| | | | | | | |
|---|---|---|---|---|---|---|
| | beta 1 | | | | | response |
| Efcab7 | EF-hand calcium binding domain 7 | 4 | 130 | 387 | 1.58 | Ion binding |
| Msh5 | mutS homolog 5 (E. coli) | 17 | 160 | 456 | 1.51 | Metabolic process |
| Stk19-ps1 | serine/threonine kinase 19, pseudogene 1 | 17 | 50.2 | 140 | 1.48 | Unknown |
| Zfp605 | zinc finger protein 605 | 5 | 263 | 655 | 1.31 | Ion binding |
| 5430403G16Rik | RIKEN cDNA 5430403G16 gene | 5 | 105 | 259 | 1.3 | Ion binding |
| RP24-201C14.6.1 | none | 3 | 61.5 | 150 | 1.29 | Unknown |
| Tnfrsf9 | tumor necrosis factor receptor superfamily, member 9 | 4 | 62.2 | 148 | 1.25 | Cell proliferation |
| H2-DMb2 | histocompatibility 2, class II, locus Mb2 | 17 | 82.1 | 190 | 1.21 | Immune response |
| Cntn1 | contactin 1 | 15 | 152 | 335 | 1.14 | Cell adhesion |
| Tmem171 | transmembrane protein 171 | 13 | 96.1 | 211 | 1.13 | Transmembrane protein |
| Cxcl10 | chemokine (C-X-C motif) ligand 10 | 5 | 104 | 227 | 1.13 | Immune response |
| Gm14038 | predicted gene 14038 | 2 | 864 | 1860 | 1.1 | pseudogene |
| Rpl21-ps10 | ribosomal protein L21, pseudogene 10 | 3 | 366 | 773 | 1.08 | ribosomal protein |
| H2-Q7 | histocompatibility 2, Q region locus 7 | 17 | 451 | 939 | 1.06 | Immune response |
| Fam129c | family with sequence similarity 129, member C | 8 | 167 | 350 | 1.06 | Unknown |
| Nnt | nicotinamide nucleotide transhydrogenase | 13 | 1480 | 3060 | 1.05 | NADP binding |
| Car3 | carbonic anhydrase 3 | 3 | 46400 | 95600 | 1.04 | Catalytic activity |
| 4930522L14Rik | RIKEN cDNA 4930522L14 gene | 5 | 322 | 649 | 1.01 | Ion binding |
| Gm5590 | predicted gene 5590 | 7 | 299 | 596 | 0.995 | pseudogene |
| Casp1 | caspase 1 | 9 | 144 | 287 | 0.994 | Catalytic activity |
| Dio3 | deiodinase, iodothyronine type III | 12 | 261 | 508 | 0.959 | Catalytic activity |
| BC064078 | cDNA sequence BC064078 | 6 | 258 | 492 | 0.933 | pseudogene |
| Tctex1d2 | Tctex1 domain containing 2 | 16 | 209 | 395 | 0.922 | Unknown |
| Gm11127 | predicted gene 11127 | 17 | 186 | 350 | 0.915 | Unknown |
| Lyz2 | lysozyme 2 | 10 | 3260 | 5980 | 0.874 | Catalytic activity |
| Cd1d1 | CD1d1 antigen | 3 | 942 | 1700 | 0.854 | Lipid metabolic process |
| Pttg1 | pituitary tumor-transforming gene 1 | 11 | 540 | 945 | 0.809 | Enzyme regulator activity |
| Rps2-ps13 | ribosomal protein S2, pseudogene 13 | X | 9060 | 15400 | 0.766 | ribosomal protein |
| Car2 | carbonic anhydrase 2 | 3 | 404 | 684 | 0.762 | Catalytic activity |
| Prune2 | prune homolog 2 (Drosophila) | 19 | 504 | 847 | 0.748 | Catalytic activity |
| Fam177a | family with sequence similarity 177, member A | 12 | 1670 | 2780 | 0.736 | Unknown |
| Gm6166 | predicted gene 6166 | 9 | 36600 | 59600 | 0.703 | pseudogene |
| Ocel1 | occludin/ELL domain | 8 | 807 | 1300 | 0.685 | Unknown |

| | | | | | | |
|---|---|---|---|---|---|---|
| | containing 1 | | | | | |
| Rps2-ps10 | ribosomal protein S2, speudogene 10 | 18 | 15500 | 24900 | 0.681 | ribosomal protein |
| Nfkbiz | nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta | 16 | 863 | 1370 | 0.67 | Immune response |
| Katnal1 | katanin p60 subunit A-like 1 | 5 | 932 | 1470 | 0.66 | Catalytic activity |
| 1700047I17Rik2 | RIKEN cDNA 1700047I17 gene 2 | 12 | 1730 | 2720 | 0.654 | Unknown |
| Gm10240 | predicted gene 10240 | 15 | 4980 | 7740 | 0.637 | Unknown |
| Lpl | lipoprotein lipase | 8 | 34700 | 53700 | 0.629 | Lipid metabolic process |
| Gm10163 | predicted pseudogene 10163 | 9 | 5170 | 7880 | 0.608 | pseudogene |
| Spon1 | spondin 1, (f-spondin) extracellular matrix protein | 7 | 4870 | 3220 | -0.597 | Cell adhesion |
| Gm5786 | predicted pseudogene 5786 | 12 | 4870 | 3150 | -0.63 | pseudogene |
| Thbs4 | thrombospondin 4 | 13 | 3690 | 2350 | -0.65 | Ion binding |
| Odz3 | odd Oz/ten-m homolog 3 (Drosophila) | 8 | 3150 | 2010 | -0.651 | Unknown |
| Gfra1 | glial cell line derived neurotrophic factor family receptor alpha 1 | 19 | 1620 | 998 | -0.696 | Cell migration |
| Fads6 | fatty acid desaturase domain family, member 6 | 11 | 825 | 501 | -0.719 | Lipid metabolic process |
| 1110005A03Rik | RIKEN cDNA 1110005A03 gene | 11 | 984 | 574 | -0.778 | Methylation |
| Glb1 | galactosidase, beta 1 | 9 | 1620 | 940 | -0.786 | Catalytic activity |
| Ppp1r35 | protein phosphatase 1, regulatory subunit 35 | 5 | 718 | 406 | -0.824 | Enzyme regulator activity |
| Rps2 | ribosomal protein S2 | 17 | 20500 | 11400 | -0.847 | ribosomal protein |
| Epha3 | Eph receptor A3 | 16 | 715 | 392 | -0.865 | Catalytic activity |
| Fabp5 | fatty acid binding protein 5, epidermal | 3 | 19500 | 10400 | -0.899 | Lipid metabolic process |
| Serpine1 | serine (or cysteine) peptidase inhibitor, clade E, member 1 | 5 | 546 | 291 | -0.905 | Cell proliferation |
| 2610035D17Rik | RIKEN cDNA 2610035D17 gene | 11 | 581 | 308 | -0.915 | lincRNA |
| Rpl35a | ribosomal protein L35A | 16 | 952 | 461 | -1.04 | ribosomal protein |
| C4b | complement component 4B (Childo blood group) | 17 | 1180 | 559 | -1.08 | Immune response |
| Gm10395 | predicted gene 10395 | 14 | 2400 | 1120 | -1.1 | Unknown |
| Entpd4 | ectonucleoside triphosphate diphosphohydrolase 4 | 14 | 3090 | 1430 | -1.11 | Metabolic process |
| D14Ertd449e | DNA segment, Chr 14, ERATO Doi 449, expressed | 14 | 2250 | 1030 | -1.13 | Unknown |
| Gm9746 | predicted gene 9746 | 14 | 2390 | 1080 | -1.14 | Unknown |
| H2-D1 | histocompatibility 2, D region locus 1 | 17 | 19800 | 8780 | -1.18 | Immune response |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gm14109 | predicted gene 14109 | 2 | 356 | 156 | -1.19 | pseudogene |
| Tmem221 | transmembrane protein 221 | 8 | 202 | 84.8 | -1.25 | Transmembrane protein |
| Gm12411 | predicted gene 12411 | 4 | 4210 | 1540 | -1.45 | pseudogene |
| Rps3a | ribosomal protein S3A | 3 | 6480 | 2370 | -1.45 | ribosomal protein |
| Glt1d1 | glycosyltransferase 1 domain containing 1 | 5 | 321 | 118 | -1.45 | Catalytic activity |
| Gm10393 | predicted gene 10393 | 14 | 1230 | 444 | -1.47 | pseudogene |
| Plac9 | placenta specific 9 | 14 | 1540 | 533 | -1.53 | Unknown |
| Gm14127 | predicted gene 14127 | 2 | 627 | 217 | -1.53 | pseudogene |
| Gm9780 | predicted gene 9780 | 14 | 1500 | 518 | -1.54 | pseudogene |
| Stfa1 | stefin A1 | 16 | 1180 | 397 | -1.58 | Enzyme regulator activity |
| Gcnt1 | glucosaminyl (N-acetyl) transferase 1, core 2 | 19 | 2240 | 739 | -1.6 | Metabolic process |
| Gm12261 | predicted gene 12261 | 11 | 255 | 77.9 | -1.71 | pseudogene |
| Cyp3a57 | cytochrome P450, family 3, subfamily a, polypeptide 57 | 5 | 181 | 54.8 | -1.72 | Metabolic process |
| Tmem79 | transmembrane protein 79 | 3 | 2330 | 695 | -1.75 | Transmembrane protein |
| B3gnt3 | UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 3 | 8 | 533 | 150 | -1.82 | Catalytic activity |
| Cpsf4l | cleavage and polyadenylation specific factor 4-like | 11 | 874 | 236 | -1.89 | Unknown |
| Ccdc122 | coiled-coil domain containing 122 | 14 | 265 | 66.7 | -1.99 | Unknown |
| Rpl21-ps8 | ribosomal protein L21, pseudogene 8 | 18 | 1480 | 338 | -2.13 | ribosomal protein |
| Gm9104 | predicted gene 9104 | 17 | 748 | 167 | -2.16 | Unknown |
| Rpl21 | ribosomal protein L21 | 5 | 2770 | 561 | -2.3 | ribosomal protein |
| 1700030C10Rik | RIKEN cDNA 1700030C10 gene | 12 | 43 | 8.16 | -2.4 | Unknown |
| Wdr95 | WD40 repeat domain 95 | 5 | 61.5 | 7.22 | -3.09 | Unknown |
| Gm11131 | predicted gene 11131 | 17 | 47.9 | 5.15 | -3.22 | lincRNA |
| Rps18 | ribosomal protein S18 | 17 | 11800 | 1040 | -3.51 | ribosomal protein |
| Rps18-ps1 | ribosomal protein S18, pseudogene 1 | 4 | 3200 | 227 | -3.82 | ribosomal protein |
| 4732490B19Rik | RIKEN cDNA 4732490B19 gene | 11 | 53.9 | 3.66 | -3.88 | lincRNA |
| Gm16867 | predicted gene, 16867 | 14 | 320 | 20.2 | -3.99 | lincRNA |
| H2-Q8 | histocompatibility 2, Q region locus 8 | 17 | 122 | 6.3 | -4.27 | Immune response |
| Rpl21-ps11 | ribosomal protein L21, pseudogene 11 | 3 | 529 | 21.4 | -4.63 | ribosomal protein |
| Gm12944 | predicted gene 12944 | 4 | 632 | 22.1 | -4.83 | pseudogene |
| Mmrn1 | multimerin 1 | 6 | 551 | 1.2 | -8.84 | Unknown |

# APPENDIX VI

## Comparative lipid analysis of *Matt*$^{ma/ma}$ and C57BL/6J epidermis

mylnefield
LIPID ANALYSIS

Email : C.Traynor@mrsltd.com Web : www.lipid.co.uk
Telephone : +44 (0)844 928 5428 Fax : +44 (0)1382 568516

Mylnefield Research Services, Errol Road, Invergowrie, Dundee DD2 5DA, UK

| | |
|---|---|
| Designation of sample - | Mouse Skin Sample<br>Code No. C57/BL6/J1 Wild Type |
| Submitted by - | University of Dundee |
| Our number - | MRS LU/2011/1961 |
| Received - | 15/11/2011 |
| Report Issue number - | 1-1 |

## TLC Report

### Extraction

The sample of mouse skin was defrosted then extracted following a method adaptation of that described by Folch *et al.* (Folch,J., Lees,M. and Stanley,G.H.S. *J. Biol. Chem.*, **226**, 497-509 (1957)) involving chloroform-methanol-potassium chloride in the ratio 2 : 1 : 0.88 v/v/v.

### 1-D Lipid Class Identification (see attached labelled scans)

A portion of mouse skin extract was dissolved in solvent and spotted onto a TLC plate along with standards for Lipid Class identification. The Plate was separated in one direction using solvent mixture 80:20:2 iso-hexane/diethyl ether/formic acid. (see attached labelled scan)
The mouse skin extract was seen to contain the following lipid classes:-
Mostly TAG with, in varying amounts CE, FFA, FC, DAG, MAG and PL.

### 2-D Phospholipid Species Identification (see attached labelled scans)

A second portion of s mouse skin extract was separated into phospholipid species using a two-dimensional solvent system. Direction 1 chloroform /methanol/water (65:25:4) followed by direction 2 chloroform/methanol/acetic acid/water (80:12:15:4).
The following species were identified Sphingomyeline, phosphatidylethanolamine, phosphatidylcholine, phosphatidylserine (slightly overlapped by PC) and phosphatidylinositol.
Two other spots seen are probably Cardiolipid (DPG) and a Lysolipid (possibly lysoPC).
Further small spots were seen on the plate which did not match our ID standards.

### 1-D Separation for Ceramide (see attached labelled scans)

A third portion of mouse skin extract was separated to show ceramide content. The extract was separated in one direction using three different solvent mixtures.
Mixture 1 dichlormethane/ethyl acetate/acetone (80:16:4)
Mixture 2 chloroform/acetone/methanol (76:8:16)
Mixture 3 isohexane/ chloroform/acetone (6:80:10:4)

Numerous bands were detected in the area where Ceramides are known to be found but we are unable to positively identify them.

mylnefield
LIPID ANALYSIS

Email : C.Traynor@mrsltd.com Web : www.lipid.co.uk
Telephone : +44 (0)844 928 5428 Fax : +44 (0)1382 568516

Mylnefield Research Services, Errol Road, Invergowrie, Dundee DD2 5DA, UK

| | |
|---|---|
| Designation of sample - | Mouse Skin Sample<br>Code No. C57/BL6/J1 Wild Type |
| Our number - | MRS LU/2011/1961 |
| Report Issue number - | 1-1 |

| | |
|---|---|
| Oil Content (%) | 22.6 |

Methods Used :- LAU_G083, LAU_G016

S. Rowbottom GRSC, CChem, MRSC
Senior Analytical Chemist — 09/12/2011

C. Traynor BSc, CChem, MRSC
Head of Mylnefield Lipid Analysis — 09/12/2011

mylnefield
LIPID ANALYSIS

Email : C.Traynor@mrsltd.com Web : www.lipid.co.uk
Telephone : +44 (0)844 928 5428 Fax : +44 (0)1382 568516

Mylnefield Research Services, Errol Road, Invergowrie, Dundee DD2 5DA, UK

| | |
|---|---|
| Designation of sample - | Mouse Skin Sample<br>Code No. MA/F6/1 Matted |
| Submitted by - | University of Dundee |
| Our number - | MRS LU/2011/1962 |
| Received - | 15/11/2011 |
| Report Issue number - | 1-1 |

## TLC Report

### Extraction

The sample of mouse skin was extracted following a method which is an adaptation of that described by Folch *et al.* (Folch,J., Lees,M. and Stanley,G.H.S. *J. Biol. Chem.*, **226**, 497-509 (1957)) involving chloroform-methanol-potassium chloride in the ratio 2 : 1 : 0.88 v/v/v.

### Lipid Class Identification (see attached labelled scans)

A portion of mouse skin extract was dissolved in solvent and spotted onto a TLC plate along with standards for Lipid Class identification. The Plate was separated in one direction using solvent mixture 80:20:2 iso-hexane/diethyl ether/formic acid.
The mouse skin extract was seen to contain the following lipid classes:-
Mostly TAG with, in varying amounts CE, FFA, FC, DAG, MAG and PL.

### 2-D Phospholipid Species Identification (see attached labelled scans)

A second portion of s mouse skin extract was separated into phospholipid species using a two-dimensional solvent system. Direction 1 chloroform /methanol/water (65:25:4) followed by direction 2 chloroform/methanol/acetic acid/water (80:12:15:4).
The following species were identified Sphingomyeline, phosphatidylethanolamine, phosphatidylcholine, phosphatidylserine (slightly overlapped by PC) and phosphatidylinositol.
Two other spots seen are probably Cardiolipid (DPG) and a Lysolipid (possibly lysoPC).
Further small spots were seen on the plate which did not match our ID standards.

### 1-D Separation for Ceramide (see attached labelled scans)

A third portion of mouse skin extract was separated to show ceramide content. The extract was separated in one direction using three different solvent mixtures.
Mixture 1 dichlormethane/ethyl acetate/acetone (80:16:4)
Mixture 2 chloroform/acetone/methanol (76:8:16)
Mixture 3 isohexane/ chloroform/acetone (6:80:10:4)

Numerous bands were detected in the area where Ceramides are known to be found but we are unable to positively identify them.

mylnefield
LIPID ANALYSIS

Email : C.Traynor@mrsltd.com Web : www.lipid.co.uk
Telephone : +44 (0)844 928 5428 Fax : +44 (0)1382 568516

Mylnefield Research Services, Errol Road, Invergowrie, Dundee DD2 5DA, UK

Designation of sample - Mouse Skin Sample
Code No. MA/F6/1 Matted

Our number - MRS LU/2011/1962

Report Issue number - 1-1

Oil Content (%) 29.7

Methods Used :- LAU_G083, LAU_G016

S. Rowbottom GRSC, CChem MRSC
Senior Analytical Chemist

S. Rowbottom 09/12/2011

C. Traynor BSc, CChem MRSC
Head of Mylnefield Lipid Analysis

C. Traynor 09/12/2011

Registered in Scotland 121376

Mylnefield Research Services Ltd. is a commercial affiliate of The James Hutton Institute.

Abbreviations:

WE - Wax Ester
CE - Cholesterol Ester
ME - Methyl Ester
TAG - Triacylglycerol
FFA - Free Fatty Acid
FC - Free Cholesterol
DAG - Diacylglycerol
MAG - Monoacylglycerol
PL - Phopholipid

Sample details

MRSLU/2011/1961 - Mouse epidermis C57BL6/J (wildtype)
MRSLU/2011/1962 - Mouse epidermis *Matt*$^{ma/ma}$ F6 (*matted*)

# Lipid Class TLC Separation
# MRSLU/2011/1961 & 1962
# solvent mix 80:20:2 isohexane/diethyl ether/formic acid

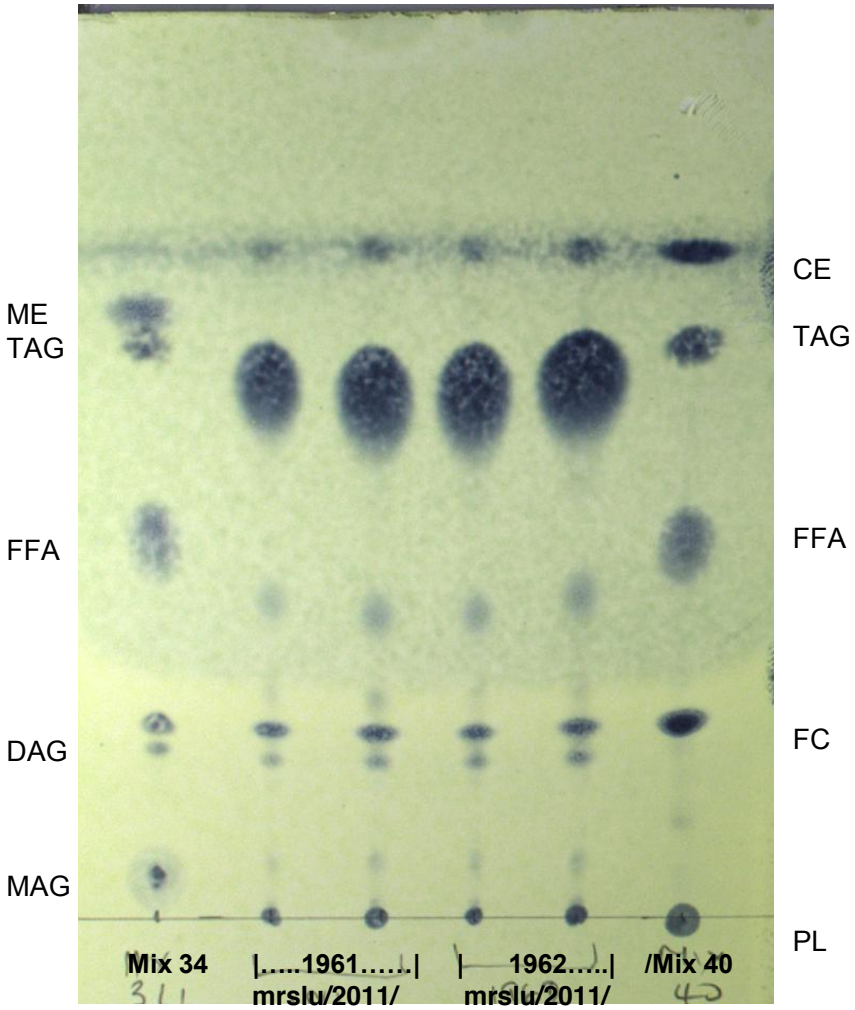



Plate sprayed with phosphomolybdic acid and charred

# Lipid Class TLC Separation
# MRSLU/2011/1961 & 1962
# solvent mix 80:20:2 isohexane/diethyl ether/formic acid

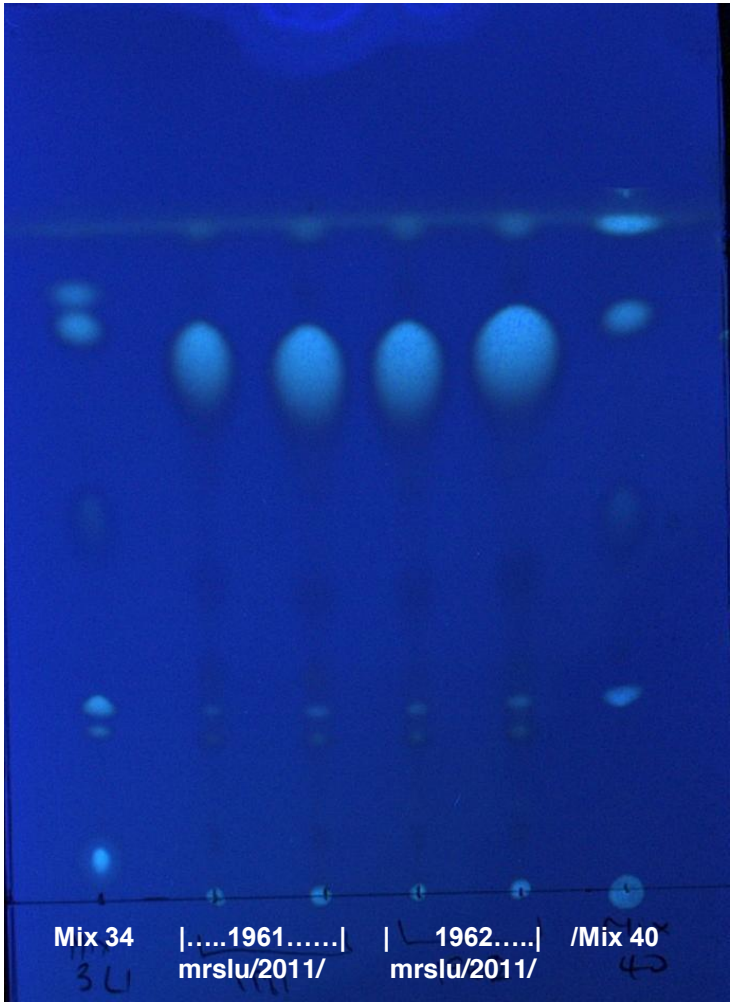



Plate sprayed with primulin and viewed under UV light

# Lipid Class 2-D TLC Separation
# MRSLU/2011/1961
# Individual Phospholipids

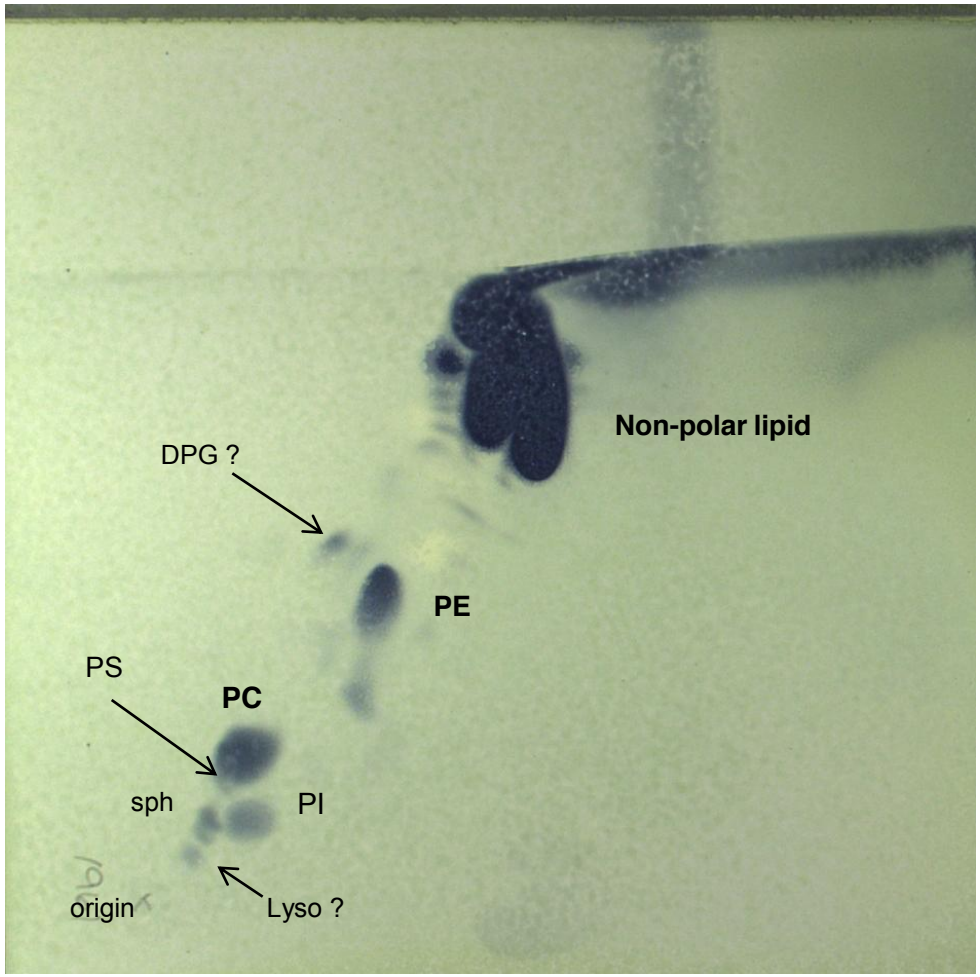



Plate sprayed with phosphomolybdic acid and charred @ 120°C

# Lipid Class 2-D TLC Separation
# MRSLU/2011/1961
# Individual Phospholipids

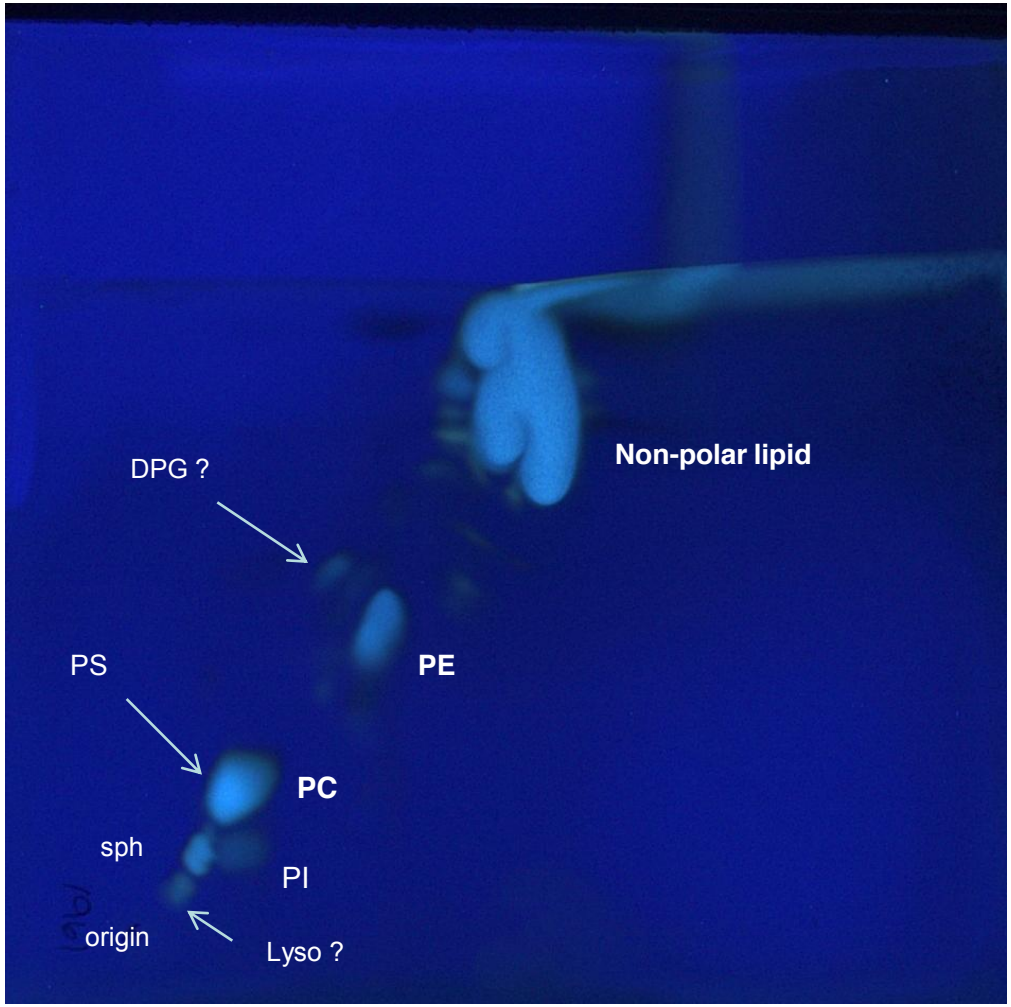



Plate sprayed with primulin and viewed under UV light

# Lipid Class 2-D TLC Separation
# MRSLU/2011/1962
# Individual Phospholipids

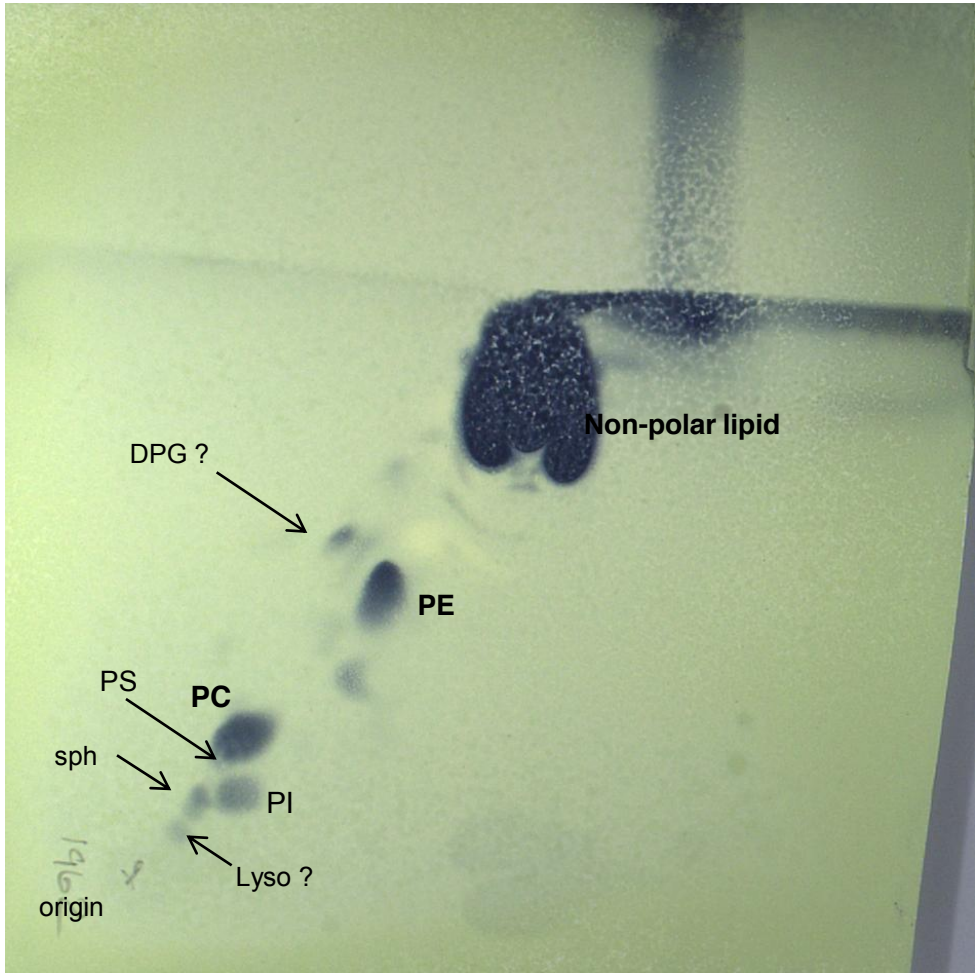



Plate sprayed with phosphomolybdic acid and charred @ 120°C

# Lipid Class 2-D TLC Separation
# MRSLU/2011/1962
# Individual Phospholipids

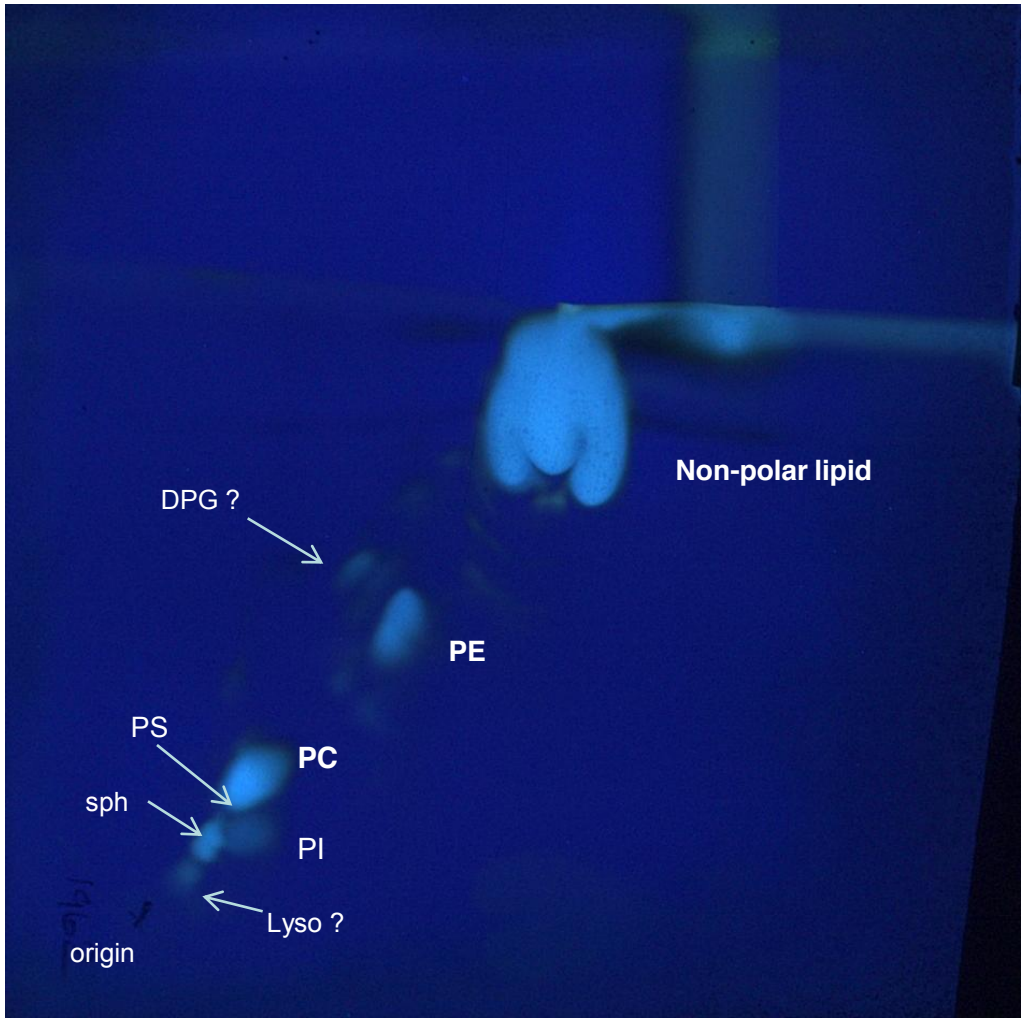



Plate sprayed with primulin and viewed under UV light

# Lipid Class 1-D TLC Separation
# MRSLU/2011/1961 & 1962
# Ceramide

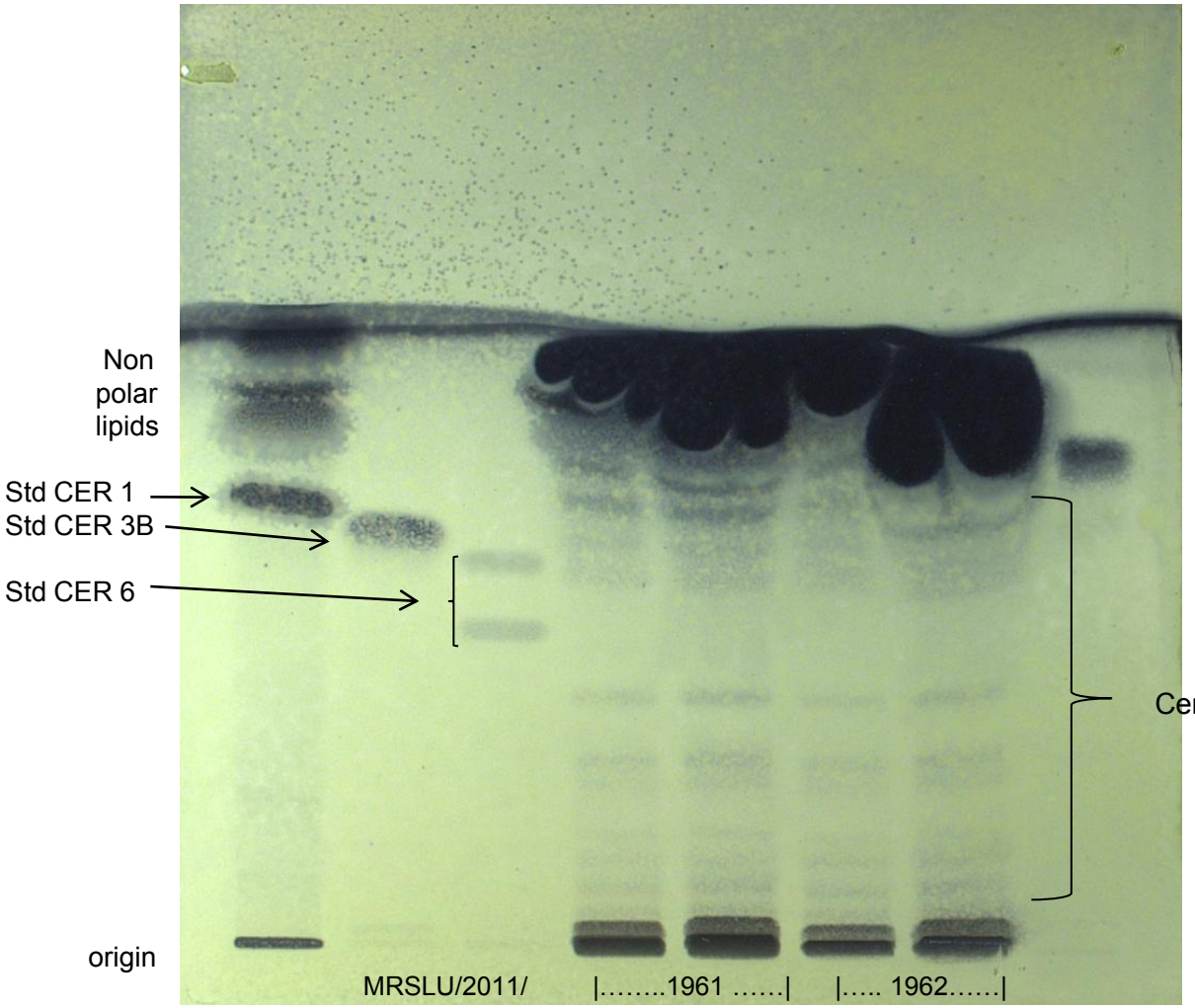



Plate sprayed with phosphomolybdic acid and charred @ 120°C

# Lipid Class 1-D TLC Separation
# MRSLU/2011/1961 & 1962
# Ceramide

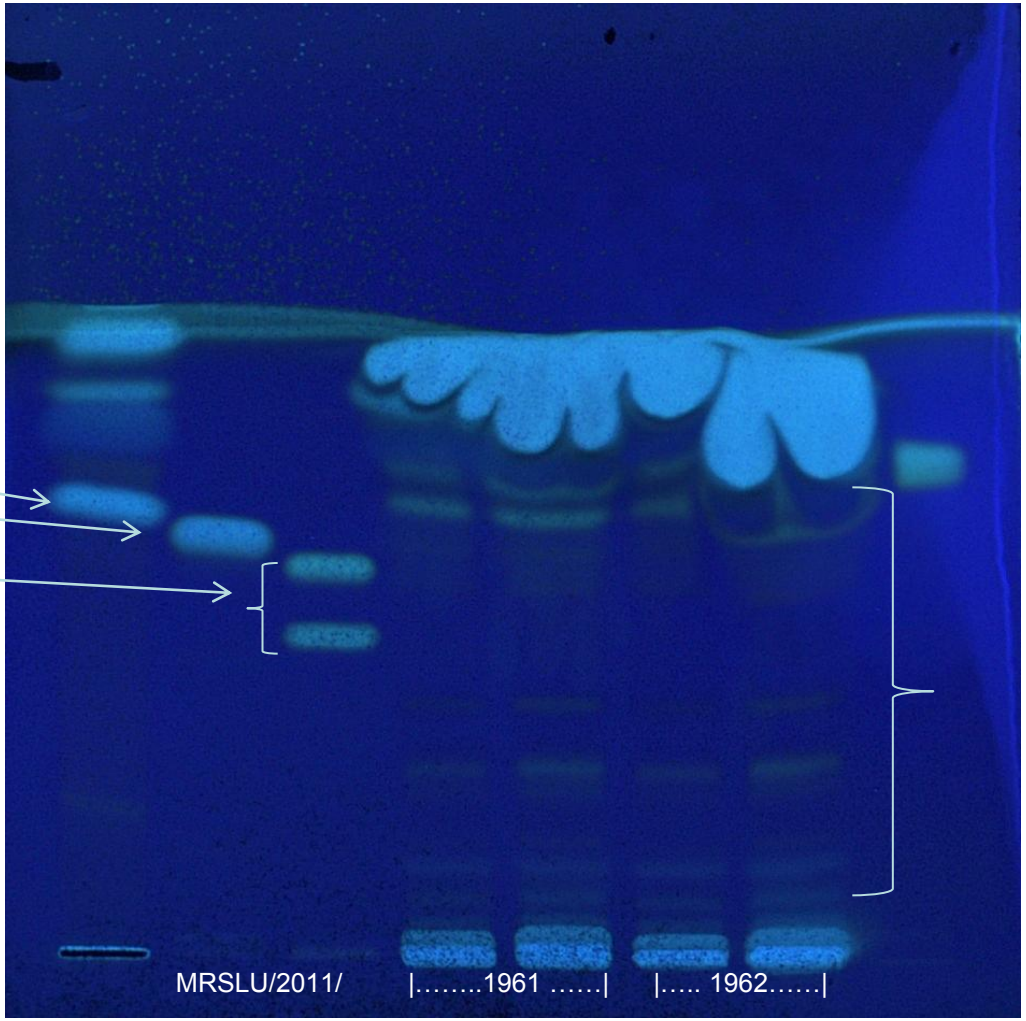



Plate sprayed with primulin and viewed under UV light

# APPENDIX VII

## Protein analysis of TMEM79

# Appendix VIII

Protein analysis of TMEM79 performed by Dr George Schneider (Bioinformatics Institute, A*STAR, Singapore).

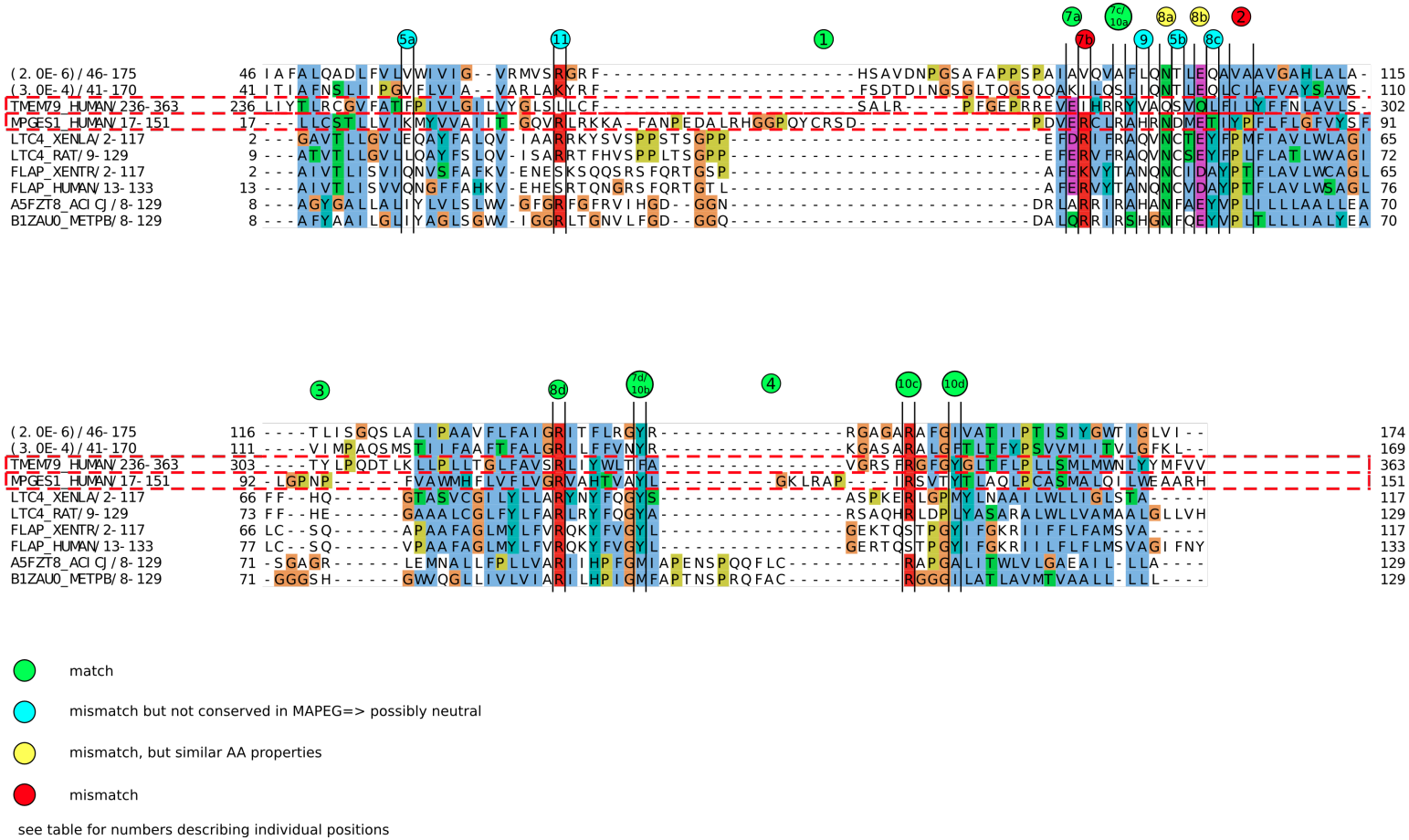



**Figure 1** Multiple protein sequence alignment of TMEM79, MPGES1, LTC4 and FLAP together with additional members of the MAPEG family. Key residues highlighted in this figure are detailed in Table 1 – 3.

**Table 1. Evidence for possible glutathione binding**

| No | MPGES1 | TMEM79 | LTC4S | FLAP | Assessment | Comment |
|---|---|---|---|---|---|---|
| 10a | Arg-70 | Arg | Arg | Thr | | The two argenine/tyrosine pairs Arg-70/Tyr-117 and Arg-126/Tyr-130 are in close contact to the glycine and cysteine moieties of GSH |
| 10b | Tyr-117 | Phe | Tyr | Tyr | | see 10a + 7d (Y117F is fully functional |
| 10c | Arg-126 | Arg | Arg | Ser | | stabilizes thiol group of GSH (also see 10a) |
| 10d | Tyr-130 | Tyr | Met/Leu | Tyr | | see 10a |
| 11 | Arg-38 | Leu | Arg | Ser | | Arg-38 makes salt-bridge to carboxylate at end of GSH (other end with Arg-70 see 10a) |

**Table 2. Overall structure**

| No | MPGES1 | TMEM79 | LTC4S | FLAP | Assessment | Comment |
|---|---|---|---|---|---|---|
| 1 | ☑ | ☑ | ☑ | ☑ | | TM1 followed by relatively large cytoplasmic loop region (highly variable) |
| 2 | Pro-81 | - | Pro (shifted) | Pro (shifted) | | conserved in MAPEG (with shift) Proline gives rise to a slight kink. |
| 3 | ☑ | ☑ | ☑ | ☑ | | short loop on lumenal side connecting TM2 and TM3 |
| 4 | ☑ | ☑ | ☑ | ☑ | | TM3 and TM4 connected by 4-residue loop |
| 5a/b | Ion-pair (Lys-26/Asp-75) | Phe... | Glu\|Leu | Gln | | Bending of TM1 through ion-pair (Lys-26/Asp-75) - not conserved |
| 6 | Arg-110 | Arg | Arg | Arg | | Arg-110 not exposed to lipid bilayer but rather involved in intermolecular contacts. |

**Table 3. Results of mutagenesis experiments on known MAPEG members**

| No | MPGES1 | TMEM79 | LTC4S | FLAP | Assessment | Comment |
|---|---|---|---|---|---|---|
| 7a | Glu-66 | Glu | Asp | Glu | | stabilization of cytoplasmic face of TM2 towards TM3 by polar side chain residues (crucial for GSH binding) |
| 7b | Arg-67 | Ile | Arg | Arg | | see 7a (mutagenesis indicates loss of function for R67A, but MGST3 has Cys and there are several Ala in the MAPEG seeds) |
| 7c | Arg-70 | Arg | Arg | Thr | | see 7a (FLAP non-catalytic and no Arg!) |
| 7d | Tyr-117 | Phe | Tyr | Tyr | | see 7a (Y117F mutant is fully functional) |
| 8a | Asn-74 | Gln | Asn | Asn | | Further stabilization toward center through side-chain interactions (Asn-74, Glu-77, Thr-78 with Arg-110, His-113). Mismatch, but also polar (there are a few sequences in the MAPEG seed that have a Gln) and Ago et al. 2007 suggest that mutation of that position in LTC4S didn't affect activity |
| 8b | Glu-77 | Gln | Asp | Asp | | see 8a.<br>Dubious since Gln is not negatively charged and E77A shows no activity, but there are 2 sequences in the MAPEG seed that have a Gln at this position. |
| 8c | Thr-78 | Leu | Tyr | Ala | | see 8a |
| 8d | Arg-110 | Arg | Arg | Arg | | see 8a (conserved in all MAPEG) |
| 8e | His-113 | Tyr | Tyr | Tyr | | see 8a |
| 9 | His-72 | Val | Gln | Asn | | His-72 forms salt bridge with Glu-77 (H72A mutant shows some activity) |

| match | mismatch but not conserved in MAPEG=> possibly neutral | mismatch, but similar AA properties | mismatch |
|---|---|---|---|

**Bioinformatics analysis of TMEM79 Protein**

## 1. Summary

A preliminary sequence analysis of TMEM79 (NP_115699.1 | CAI14162.1) was performed using ANNOTATOR analysis framework (Schneider et al. 2010; Ooi et al. 2009). There seems to be evidence from computational predictions and the comparison to published structures to support the following findings:

1. TMEM79 and its orthologs are distantly related to the MAPEG (Membrane Associated Proteins in Eicosanoid and Glutathione metabolism) superfamily many of whose members are involved in inflammation pathways.

2. Although sequence identity is very low, a multiple sequence alignment of MPGES1 and TMEM79 together with other canonical members of the MAPEG family shows that a large percentage of the key residues involved in glutathione binding in MPGES1 are also present in TMEM79 (see section 3).

3. Since MAPEG family members catalyze glutathione-dependent transformations of lipophilic substrates harvested from the lipid bilayer (Martinez Molina et al. 2008), experimentally establishing Glutathione binding of TMEM79 could serve as a first step to validate the above predictions and decide on whether to continue with a search for a lipophilic substrate (see section 4).

## 2. General Sequence Analysis

A segment based strategy for functional analysis was used. Around 30 algorithms were applied to try to distinguish non-globular segments from globular domains. The non-globular segments are often regions of low-sequence complexity or compositional bias, but may carry out important structural functions or contain signals for posttranslational modification or localization. On the other hand, conserved globular domains with a fairly stable three-dimensional structure might carry out enzymatic function or serve as an interface in protein-protein interactions or have general scaffolding roles. The results can be seen in (Figure 1) and show that TMEM79 has a long N-terminal region of low sequence complexity. No further analysis of this region was undertaken. The C-terminal half of the protein is dominated by five predicted trans-membrane segments. After this sequence motif and architecture analysis, orthologous sequences were gathered which were subjected to a similar analysis as above to look for emerging patterns. Subsequently, the analysis was extended to individual segments to not only include orthologs but also remote homologs, which were gathered with an iterative procedure.

The ortholog collection algorithm was used to find a number of orthologs and subsequent analysis of these found an enrichment for PFAM domain hits (PF01124) against the MAPEG protein family

Further evidence for membership in the MAPEG family was uncovered:

1. Inspection of the HMMER result on TMEM79 finds a sub-significant hit (E-value: 0.029) against the MAPEG domain (PF01124)

2. PSI-BLAST finds members of the MAPEG family starting from round 2.

3. The locally installed version of HHPRED (Soding et al. 2005) finds a clear hit (E-value: 9.7E-11) against the MAPEG domain (PF01124)

A number of caveats should be mentioned:

- All the well-studied MAPEG members only have four TM segments, while TMEM79 has five. The region of homology clearly only extends to the last four segments of TMEM79.

- All MAPEG members appear to only as long as the TM segments together. There are no long N-terminal unstructured regions, as found in TMEM79/mattrin.

- All known structures of MAPEG members show the formation of a homotrimer.

## 3. Comparison of TMEM79 and MPGES1

There are a number of publications that provide detailed structural information as well as results of site-directed mutagenesis for MPGES1 (Microsomal prostaglandin E synthase 1) (Jegerschöld et al. 2008), LTC4S (Leukotriene C4 synthase) (Molina et al. 2007; Ago et al. 2007) and FLAP (5-Lipoxygenase-activating protein) (Ferguson et al. 2007) with the last one lacking enzymatic activity.

A multiple sequence alignment was constructed using locally installed MAFFT (Katoh et al. 2005) and taking as input the orthologs of TMEM79, MPGES1 and the seed sequences for the MAPEG domain in PFAM (PF01124). Since the seed only contained xenopus sequences for LTC4S and FLAP, the rat and human orthologs were added respectively. A small number of MAPEG seed sequences with very long loop regions were removed.

The alignment shown in (Figure 1), above, presents alignment with only the core members shown. The first two sequences are orthologs (YP_617596.1 and ZP_05250079.1) of TMEM79, which were retained to provide some visual guidance and context. The third and fourth sequences are TMEM97 and MPGES1 respectively (they are surrounded by a red dotted box). They are followed by two sequences of LTC4 and FLAP each. The last two sequences belong to the MAPEG seed and are again left there to provide some context.

Tables 1 - 3 contain assessment which was done by comparing features and individual residues mentioned in (Molina et al. 2007) among the different sequences. Mismatches between residues are classified and color-coded as explained below. The same numbers and color-codes are indicated in the multiple

sequence alignment shown in (Figure 1).

## 4. Summary

In summary, there seems to be a sufficient match between key residues TMEM79 of glutathione binding to merit further experimental investigation as a MAPEG-like protein.

## 5. Bibliography


Ago, H. et al., 2007. Crystal structure of a human membrane protein involved in cysteinyl leukotriene biosynthesis. Nature, 448(7153), pp.609-612.

Ferguson, A.D. et al., 2007. Crystal structure of inhibitor-bound human 5-lipoxygenase-activating protein. Science (New York, N.Y.), 317(5837), pp.510-512.

Jegerschöld, C. et al., 2008. Structural basis for induced formation of the inflammatory mediator prostaglandin E2. Proceedings of the National Academy of Sciences, 105(32), pp.11110 -11115.

Katoh, K. et al., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic acids research, 33(2), pp.511–518.

Martinez Molina, D., Eshaghi, S. & Nordlund, Pär, 2008. Catalysis within the lipid bilayer-structure and mechanism of the MAPEG family of integral membrane proteins. Current Opinion in Structural Biology, 18(4), pp.442-449.

Molina, D.M. et al., 2007. Structural basis for synthesis of inflammatory mediators by human leukotriene C4 synthase. Nature, 448(7153) pp.613-616.

Ooi, H.S. et al., 2009. ANNIE: integrated de novo protein sequence annotation. 37 (Web Server issue), p.W435-W440.

Schneider, G. et al., 2010. Integrated tools for biomolecular sequence-based function prediction as exemplified by the ANNOTATOR software environment. Methods in Molecular Biology (Clifton, N.J.), 609, pp.257-267.

Soding, J., Biegert, A. & Lupas, A.N., 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Research, 33(Web Server), p.W244-W248.

# APPENDIX VIII

## Case-control studies for meta-analysis of atopic dermatitis (AD) patients from England, UK, Ireland, Germany and Scotland

## APPENDIX VIII

**Case-control studies for meta-analysis of atopic dermatitis patients from England, UK, Ireland, Germany and Scotland.**

**Authors involved:** Christabelle S. M. Goh[1], Stephan Weidinger[2], Hansjörg Baurecht[2], Young-Ae Lee[3], Jonathan N. W. N. Barker[4], Nick J. Reynolds[5], Heather J. Cordell[5], Sara J. Brown[1], Alan D. Irvine[6], W.H. Irwin McLean[1], Aileen Sandilands[1]

*1. University of Dundee, United Kingdom; 2. University of Kiel, Germany; 3. Charite, Berlin; 4. King's College London, United Kingdom; 5. Newcastle University, United Kingdom; 6. Trinity College Dublin, Ireland*

Five independently recruited AD case collections were compared with ethnically matched population controls. Statistical and meta-analysis was conducted by Dr. Sara J. Brown (University of Dundee).

(1) English adult severe AD is early onset, persistent and severe disease, diagnosed by experienced dermatologists in secondary and tertiary care.

(2) UK mild-moderate pediatric AD includes cases collected from an English population birth cohort (n=177) and a Scottish Primary Care collection (n=161). AD is defined using the UK refinement of the Hanifin and Rajka diagnostic criteria.

(3) Irish pediatric AD cases were collected in secondary and tertiary care clinics in Ireland, defined by experienced dermatologists.

(4) German AD cases are adults and children diagnosed by experienced dermatologists and pediatricians using standard criteria for AD, in tertiary care clinics.

(5) Scottish asthma cases with AD are individuals with physician-diagnosed asthma in primary and secondary care, with parent- or self-reported history of ever having had AD.

The following matched control populations were used:

(1) 1958 Birth Cohort comprises unselected English population controls.

(2) English pediatric controls were ascertained not to have AD at the age of 7-9 years. They are the hyper-normal controls from the same English population birth cohort from which the English mild-moderate pediatric AD cases were recruited.

(3) Irish population controls are healthy adult blood donors in Dublin, Ireland.

(4) German population controls include pediatric and adult cohorts from Germany.

(5) Scottish population controls are healthy adult blood donors from throughout Scotland.

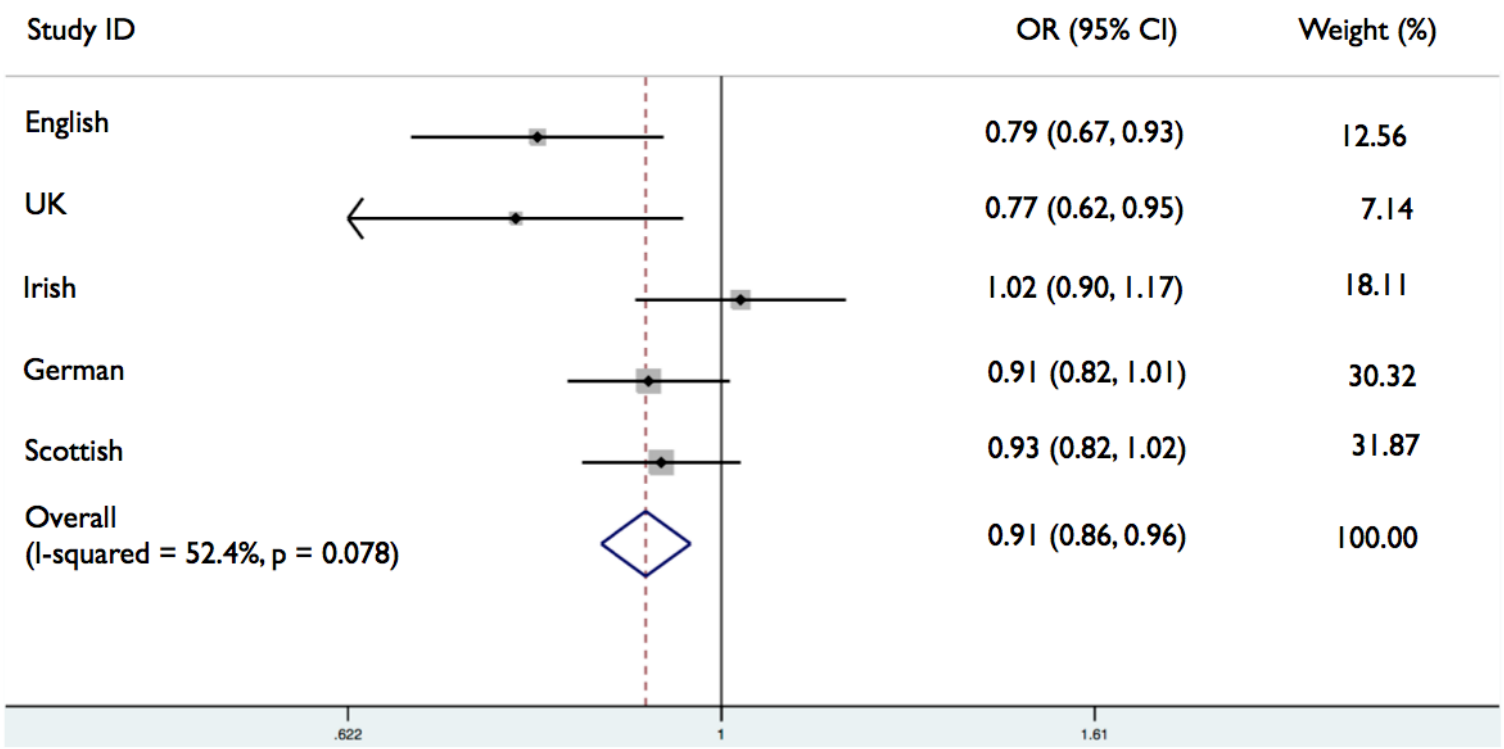



**Figure 1** Forrest plot showing results of meta-analysis of five case-control studies to investigate the association of rs6684514 and AD. Study populations used were: English, English adult severe AD vs English population controls from the 1958 Birth Cohort; UK; UK, UK mild-moderate pediatric AD vs English pediatric controls without AD; Irish, Irish pediatric AD vs Irish adult population controls; German, German AD cases vs German population controls; Scottish; Scottish asthma cases with AD vs Scottish population controls. ES = estimated odds ratio; CI = confidence interval. Meta-analysis carried out using the "metan" function in Stata® (StataCorp, College Station, Texas).

Case-control analyses of two independent AD case collections, English adult AD and UK pediatric AD, with separate population-matched controls showed a significant association between rs6684514 and AD. The minor allele (A) is protective for disease in both English adult AD (OR ~0.791, p=0.0038) and UK pediatric AD (OR ~0.770, p=0.0143). Controlling for the association of *FLG* null mutations in these analyses demonstrated that the association of rs6684514 with AD is independent of the strong and significant *FLG* association. The significant association was further replicated in a German AD case-control analysis in which the OR after controlling for *FLG* mutations was 0.86, 95%CI 0.76 to 0.97, p=0.0161. However, the association of rs6684514 was not replicated in the Irish case/control analyses (OR ~1.02, p=0.129).

Meta-analysis of all 4,245 AD cases from England, Scotland, Ireland and Germany with 10,558 population-matched controls using a fixed-effects model did not show significant heterogeneity (p=0.078). The small but significant effect of rs6684514 on AD risk was confirmed in the meta-analysis, OR 0.91, 95% CI 0.86 to 0.96, p=0.078 (**Figure 1**).

# APPENDIX IX

## Publications

# *Tmem79/Matt* is the matted mouse gene and is a predisposing gene for atopic dermatitis in human subjects

**Sean P. Saunders, PhD,[a,b,c]* Christabelle S. M. Goh, BSc,[d]* Sara J. Brown, MD,[d] Colin N. A. Palmer, PhD,[e] Rebecca M. Porter, PhD,[a,d] Christian Cole, PhD,[d,f] Linda E. Campbell, MSc,[d] Marek Gierlinski, PhD,[f] Geoffrey J. Barton, PhD,[f] Georg Schneider, PhD,[g] Allan Balmain, PhD,[h] Alan R. Prescott, PhD,[i] Stephan Weidinger, MD, PhD,[j] Hansjörg Baurecht, PhD,[j] Michael Kabesch, MD,[k] Christian Gieger, PhD,[l] Young-Ae Lee, MD,[m] Roger Tavendale, PhD,[e] Somnath Mukhopadhyay, MD,[n] Stephen W. Turner, MD,[o] Vishnu B. Madhok, MD,[p] Frank M. Sullivan, MD,[p] Caroline Relton, PhD,[q] John Burn, MD,[q] Simon Meggitt, MD,[r] Catherine H. Smith, MD,[s] Michael A. Allen, PhD,[s] Jonathan N. W. N. Barker, MD,[s] Nick J. Reynolds, MD,[t] Heather J. Cordell, DPhil,[q] Alan D. Irvine, MD,[a,b,u] W. H. Irwin McLean, DSc,[d] Aileen Sandilands, PhD,[d] and Padraic G. Fallon, PhD[a,b,c]** *Dublin, Ireland, Dundee, Brighton, Aberdeen, Newcastle, Newcastle-upon-Tyne, and London, United Kingdom, Singapore, San Francisco, Calif, and Kiel, Regensburg, Neuhersberg, and Berlin, Germany*

**Background: Atopic dermatitis (AD) is a major inflammatory condition of the skin caused by inherited skin barrier deficiency, with mutations in the filaggrin gene predisposing to development of AD. Support for barrier deficiency initiating AD came from flaky tail mice, which have a frameshift mutation in *Flg* and also carry an unknown gene, *matted*, causing a matted hair phenotype.**

**Objective: We sought to identify the *matted* mutant gene in mice and further define whether mutations in the human gene were associated with AD.**

**Methods: A mouse genetics approach was used to separate the *matted* and *Flg* mutations to produce congenic single-mutant strains for genetic and immunologic analysis. Next-generation sequencing was used to identify the *matted* gene. Five**

From [a]the School of Medicine, Trinity College Dublin; [b]the National Children's Research Centre, Our Lady's Children's Hospital, Dublin; [c]Trinity Biomedical Sciences Institute, Trinity College Dublin; [d]the Centre for Dermatology and Genetic Medicine, [e]the Biomedical Research Institute, [f]the Bioinformatics Research Group, [i]the Division of Cell Signalling and Immunology, and [p]Population Health Sciences, University of Dundee; [g]the Bioinformatics Institute, A*STAR, Singapore; [h]the Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco; [j]the Department of Dermatology, Allergy, and Venerology, University Hospital Schleswig-Holstein, Kiel; [k]the Department of Pediatric Pneumology and Allergy, Children's Hospital Regensburg; [l]the Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg; [m]Pediatric Pneumology and Immunology, Charité University Medicine Berlin and Max Delbrück Center for Molecular Medicine (MDC) Berlin-Buch, Berlin; [n]the Academic Department of Paediatrics, Royal Alexandra Children's Hospital, Brighton and Sussex Medical School, Brighton; [o]the Department of Child Health, University of Aberdeen; [q]the Institute of Genetic Medicine, Newcastle University; [r]the Department of Dermatology, Royal Victoria Infirmary, Newcastle-upon-Tyne; [s]St John's Institute of Dermatology, King's College London; [t]the Institute of Cellular Medicine, Newcastle University; and [u]the Department of Paediatric Dermatology, Our Lady's Children's Hospital, Crumlin, Dublin.

*These authors contributed equally to this work.

Supported by the Wellcome Trust (Programme grant 092530/Z/10/Z to W.H.I.M., A.D.I., G.J.B., and P.G.F.; Bioresources grant 090066/B/09/Z to A.D.I., W.H.I.M., and S.J.B.; Clinical Intermediate Fellowship WT086398MA to S.J.B.; and Senior Fellowship 087436 to H.J.C.); the Interdisciplinary Training Programme for Clinicians in Translational Medicine and Therapeutics (to N.J.R.); Wellcome Trust Bioresources grant 099177/Z/12/Z (to C.N.A.P.); the Medical Research Council (Programme grant G0802780 to W.H.I.M.); the Science Foundation Ireland (to P.G.F.); the National Children's Research Centre (to P.G.F. and A.D.I.); and the National Cancer Institute (NCI U01 CA141455 to A.B.). The Centre for Dermatology and Genetic Medicine, University of Dundee, is supported by a Wellcome Trust Strategic Award (098439/Z/12/Z to W.H.I.M. and G.J.B.). The PAGES, BREATHE, and GS:3D cohorts were funded by the Scottish Executive Health Department (CZB/4/285).

Disclosure of potential conflict of interest: S. P. Saunders has been supported by one or more grants from the Wellcome Trust. S. J. Brown has been supported by one or more grants from the Wellcome Trust, National Children's Research Centre Dublin (Ireland), and Anonymous Trustees, Dundee, United Kingdom; is employed by NHS Tayside; and has received one or more payments for lecturing from or is on the speakers' bureau for the American Academy of Allergy, Asthma & Immunology. C. Cole has been supported by one or more grants from the Wellcome Trust. G. J. Barton has been supported by one or more grants from the Wellcome Trust. A. Balmain has been supported by one or more grants from the National Cancer Institute. S. Weidinger has been supported by one or more grants from the National Genome Research Network. M. Kabesch has been supported by one or more grants from the European Union, the German Ministry of Education and Research, and the German Research Foundation and has received one or more payments for lecturing from or is on the speakers' bureau for ERS, EIAACI, ATS, Novartis, and GlaxoSmithKline. Y.-A. Lee has been supported by one or more grants from the German Ministry of Education and Research and from the German Research Foundation. S. Mukhopadhyay has been supported by one or more grants from Scottish Enterprises, the Perth and Kinross Council, the Gannochy Trust, and the Translational Medical Research Consortium; has received one or more consulting fees or honoraria from the Translational Medicine Research Consortium; is employed by Brighton and Sussex Medical School; has provided expert testimony for the Central Legal Office, Scotland; has received one or more grants from or has one or more grants pending with the National Institute of Health Research; has one or more patents planned, pending, or issued unrelated to this article; and has received one or more payments for the development of educational presentations for Thermo Fisher. F. M. Sullivan has been supported by one or more grants from the Chief Scientist Office Scotland. J. Burn is the Medical Director for Quantum Dx Ltd. H. J. Cordell is a Board member for Wiley-Blackwell Life Science Journals; is employed by Newcastle University; has received one or more grants from or has one or more grants pending with the Wellcome Trust, the Medical Research Council, and NIAAA; and has received one or more payments for lecturing from or is on the speakers' bureau for Erasmus University Rotterdam, the Wellcome Trust, Rockefeller University, and the Ontario Institute for Cancer Research. N. J. Reynolds has been supported by a Wellcome Trust Research Leave Award for Clinical Academics. A. D. Irvine has been supported by one or more grants from the Wellcome Trust. P. G. Fallon has been supported by one or more grants from Science Foundation Ireland, the National Children's Research Centre, and the Wellcome Trust. The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication June 24, 2013; revised August 15, 2013; accepted for publication August 16, 2013.

Corresponding author: Padraic G. Fallon, PhD, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin 2, Ireland. E-mail: pfallon@tcd.ie. Or: W. H. Irwin McLean, DSc, Centre for Dermatology and Genetic Medicine, University of Dundee, Dundee, United Kingdom. E-mail: w.h.i.mclean@dundee.ac.uk.

0091-6749/$36.00



http://dx.doi.org/10.1016/j.jaci.2013.08.046

**independently recruited AD case collections were analyzed to define associations between single nucleotide polymorphisms (SNPs) in the human gene and AD.**

**Results: The matted phenotype in flaky tail mice is due to a mutation in the *Tmem79/Matt* gene, with no expression of the encoded protein mattrin in the skin of mutant mice. *Matt*$^{ft}$ mice spontaneously have dermatitis and atopy caused by a defective skin barrier, with mutant mice having systemic sensitization after cutaneous challenge with house dust mite allergens. Meta-analysis of 4,245 AD cases and 10,558 population-matched control subjects showed that a missense SNP, rs6694514, in the human *MATT* gene has a small but significant association with AD.**

**Conclusion: In mice mutations in *Matt* cause a defective skin barrier and spontaneous dermatitis and atopy. A common SNP in *MATT* has an association with AD in human subjects.**

**(J Allergy Clin Immunol 2013;■■■:■■■-■■■.)**

***Key words:*** *Allergy, association, atopic dermatitis, atopy, eczema, filaggrin, flaky tail,* Matt*, mattrin, mouse, mutation,* Tmem79

Atopic dermatitis (AD) is the most common diagnosis in dermatology, affecting approximately 1 in 5 children in the developed world,[1] and is frequently associated with atopic asthma and a wide range of allergies.[2] AD is a highly heritable complex trait; however, environmental influences also play a role in triggering the atopic diathesis.[3] Genome-wide association studies in AD have identified several susceptibility loci[4-7]; however, the major and only functionally characterized genetic factor is the filaggrin gene *(FLG)*, which encodes the skin barrier protein filaggrin.[3] Prevalent loss-of-function variants in *FLG* were identified as the cause of the single-gene disorder ichthyosis vulgaris (dry flaky skin).[8] Soon thereafter, these variants were shown to be strongly associated with AD,[9] with heterozygous odds ratios (ORs) of greater than 7 and homozygous ORs of greater than 150 in case-control studies in which both prevalent and rare variants were analyzed.[10]

The hypothesis that skin barrier deficiency in the context of *FLG* mutations is an initiator of AD was confirmed experimentally by using the flaky tail mouse mutant,[11] which was shown to carry a frameshift mutation in the murine *Flg* gene.[12] Flaky tail mice have a defective skin barrier, with increased percutaneous transfer of antigens and chemical haptens.[12-15] The *ft* mutation arose spontaneously in 1958 in the progeny of crosses between heterogeneous stocks of mice with the recessive mutation *matted (ma)*, and these mutations are maintained as a double-mutant (DM) strain known as *maft*. The matted hair phenotype was used for many years as a surrogate marker for the *ft* mutation because, remarkably, the *ft* and *ma* mutations are closely linked on chromosome 3 in the mouse.[16] The DM *maft* mice have been routinely used for studies of skin barrier–deficient AD in recent years.[13-15,17] In mice and human subjects the *FLG* gene resides in the epidermal differentiation complex, a cluster of more than 70 genes encoding proteins involved in skin barrier formation and differentiation of stratified epithelia,[18] including those within the hair follicle.[18-21] We suspected the nearby *ma* gene might also be involved in epithelial barrier function, and in this study we set out to separate this allele from *Flg*$^{ft}$ and identify the causative defect.

*Abbreviations used*

| | |
|---|---|
| AD: | Atopic dermatitis |
| DM: | Double mutant |
| FLG: | Filaggrin |
| HDM: | House dust mite |
| hpf: | High-power field |
| MAPEG: | Membrane-associated proteins in eicosanoid and glutathione metabolism |
| OR: | Odds ratio |
| SNP: | Single nucleotide polymorphism |
| TEWL: | Transepidermal water loss |
| WT: | Wild-type |

## METHODS

### Isolation of the matted mouse strain

DM *Matt*$^{ma/ma}$*Flg*$^{ft/ft}$ mice were provided by Dr John P. Sundberg (Jackson Laboratory, Bar Harbor, Me).[12] DM mice were crossed with C57BL/6J mice to generate *Matt*$^{ma/+}$*Flg*$^{ft/+}$ mice. The *Flg*$^{ft}$ and *Matt*$^{ma}$ mutations were separated and backcrossed to congenic C57BL/6J background in accordance with the breeding strategy outlined (see Fig E1 in this article's Online Repository at www.jacionline.org). C57BL/6J mice were used as wild-type (WT) control animals. B6.CBAGr-ma/J (JAX *Matt*$^{ma/ma}$) mice were obtained from the Jackson Laboratory. Mice were housed in specific pathogen-free conditions, with irradiated diet and bedding and water *ad libitum*. All animal experiments were performed in compliance with Irish Department of Health and Children regulations and approved by Trinity College Dublin's BioResources Ethical Review Board.

### Gene mapping

Skin samples were obtained from neonatal mice, and DNA was extracted by using the DNA Purification Kit (Promega, Madison, Wis). Genomic DNA extracted from murine neonatal blood was amplified with the GoTaq Flexi DNA Polymerase kit (Promega). All samples were sequenced by using the ABI 3730 DNA Systems (Applied Biosystems, Foster City, Calif). Mapping primers were used to amplify and sequence murine chromosome 3 (see Table E1, *A*, in this article's Online Repository at www.jacionline.org). *Matted* genomic sequences were compared with C57BL/6J for regions of congenicity.

### Next-generation sequencing and bioinformatics

Three replicates from each sample (WT and *Matt*$^{ma/ma}$) were submitted for next-generation sequencing. The replicates were run multiplexed on an Illumina GAIIx and HiSeq 2000 (Illumina, San Diego, Calif) by using v3 sequencing chemistry and the Roche 454 Titanium workflow (Roche, Mannheim, Germany). For further details on sequencing, bioinformatics, and single nucleotide polymorphism (SNP) and InDel methodologies, see the Methods section in this article's Online Repository at www.jacionline.org.

### Analysis and identification of murine *Tmem79/Matt* gene

Each of the 4 exons was amplified individually by using PCR with the following conditions for all exons: 1 cycle of 94°C for 5 minutes; 35 cycles of 94°C for 30 seconds, 54°C for 30 seconds, and 72°C for 1 minute; and a final extension at 72°C for 5 minutes.

### Semiquantitative RT-PCR

Mouse tissue samples were lysed with TissueLyser LT (Qiagen, Hilden, Germany), and RNA was extracted with the RNeasy kit (Qiagen). RNA was reverse transcribed with the ImProm-II Reverse Transcription System (Promega). An intron-spanning amplification was carried out on *Tmem79/Matt* across exons 3 and 4. *Krt14* was used as a loading control. RT-PCR primers used are shown in Table E1, *B*, in this article's Online Repository.

## Mouse genotyping

The 644-bp PCR product of the third *Tmem79/Matt* exon was digested with the restriction enzyme CviQI (New England Biolabs, Ipswich, Mass), and digested fragments were separated on agarose gel by using electrophoresis. Primers used are shown in Table E1, *A*. The $Flg^{ft}$ mutation was genotyped, as previously described.[12]

## Immunoblotting

The epidermis was separated from the dermis of neonatal mice after immersion in 5 mmol/L EDTA at 50°C for 5 minutes, followed by cooling in ice-cold PBS. The separated epidermis was extracted in urea/Tris buffer containing protease inhibitor cocktail (Halt; Thermo Scientific, Erembodegem, Belguim) by means of homogenization. Protein samples were separated on SDS-polyacrylamide gels and transferred to polyvinylidene difluoride membranes (Millipore, Temecula, Calif), which were probed with rabbit polyclonal antibodies specific for human mattrin (TMEM79; Novus Biologicals, Littleton, Colo). Primary antibodies were detected by means of incubation with a horseradish peroxidase–conjugated goat anti-rabbit secondary antibody (Dako, Stockport, United Kingdom). Immunolabeled proteins were visualized by using chemiluminescence with the ECL detection system (Millipore).

## Immunofluorescence and Nile Red staining

Dorsal murine sections were obtained from 4-day-old neonatal mice and snap-frozen immediately. All frozen samples were then cryosectioned at 4 to 5 μm and stored at 80°C until use. After drying, sections were fixed in 50% methanol acetone, washed in PBS, and incubated with primary antibodies. Nuclei were stained with 4′-6-diamidino-2-phenylindole (Sigma-Aldrich, St Louis, Mo), and slides were fixed in Hydromount (National Diagnostics, Atlanta, Ga). Primary antibodies used in whole mounts were as follows: polyclonal TMEM79 (TMEM79; Novus Biologicals) and monoclonal CK-13 against cytokeratin 17 (Sigma-Aldrich). Frozen mouse skin sections were stained with Nile Red and 6-diamidino-2-phenylindole. Images were collected with a Zeiss LSM700 confocal microscope (Zeiss, Oberkochen, Germany).

## Quantitative PCR of human cDNA tissue array

Expression of *TMEM79/MATT* cDNA was quantified across a multiple-tissue panel containing cDNA from 48 different human tissues (Origene Technologies, Rockville, Md). This was done with the TaqMan Gene Expression Assay, C_25986870_10 (Life Technologies, Carlsbad, Calif), which spans exons 3 to 4 to result in amplicon length of 65 bp in cDNA sequence. The standard used in this assay was a commercially available plasmid containing the full-length WT *TMEM79/MATT* (Origene Technologies). A standard curve was calculated by using the formulae provided by Applied Biosystems.

## Phenotypic scoring

The severity of inflammation and AD-like pathology was scored by using the macroscopic diagnostic criteria described for the skin inflammation seen in the Nc/Nga mouse model of AD.[22] Briefly, a scoring system (0, none; 1, mild; 2, moderate; and 3, severe) was applied to the symptoms of pruritus, edema, erosion, scaling, and erythema. Pruritus (scratch behavior) was observed for 5 minutes. Total scores for each mouse were calculated from the sum of individual scores.

## Cutaneous challenge with house dust mite antigen

Mice were challenged with house dust mite (HDM) to induce allergic skin sensitization, as previously described.[13] Briefly, HDM (Greer Laboratories, Lenoir, NC) was topically applied to intact skin of WT and $Matt^{ma}$ mice by using Finn chamber patch tests (SmartPractice, Phoenix, Ariz) and secured with Scanpor tape (Bio-Diagnostics, Worcester, United Kingdom). The regimen involved five 1-week exposures to the allergen patch separated by 2-week intervals. Hair was clipped 24 hours before application of HDM that was prepared in endotoxin-free Dulbecco PBS and vehicle to a concentration of 1 mg/mL. Dulbecco PBS was used as the control. Mice were assessed 24 hours after the final HDM challenge.

## ELISA

Serum levels of HDM-specific IgE, $IgG_1$, and $IgG_{2a}$ were detected by means of direct ELISA. Total serum IgE levels were measured with a sandwich ELISA, according to the manufacturer's instructions (BD PharMingen, San Jose, Calif).

## Histology

Dorsal skin sections were removed and fixed in 10% formal saline. Paraffin-embedded sections were stained with hematoxylin and eosin and examined in blinded fashion by 2 observers independently. To quantify dermal cell numbers, we used a previously described scoring system.[12] The total number of cells per high-power field (hpf) of view were counted on 20 hpfs per mouse. For skin sections, an arbitrary histologic scoring system was used to quantify acanthosis and hyperkeratosis. Acanthosis scoring was based on the magnitude of epidermal hyperplasia, and hyperkeratosis scoring was based on the magnitude of stratum corneum thickening, with scores (0, basal; 1, mild; 2, moderate; 3, marked; 4, very marked; and 5, extreme pathology) for each parameter based on measurements per hpf in sections.

## Measurement of transepidermal water loss

A Courage and Khazaka Tewameter TM210 (Enviroderm, Evesham, United Kingdom) was used for measurement of transepidermal water loss (TEWL), which was measured on the dorsal flank of mice 24 hours after hair clipping. TEWL was recorded at an ambient temperature of 19°C to 21°C and humidity of 50% ± 5%. In mice sensitized to HDM, TEWL was measured in PBS- and HDM-treated mice before and after the HDM challenge regimen.

## Electron microscopy

For scanning electron microscopy, hair fibers were removed from mice and fixed in cacodylate-buffered glutaraldehyde. Samples were prepared for scanning electron microscopy by using established procedures[11] and examined in a Zeiss Supra scanning electron microscope.

## Statistical analysis of mouse studies

GraphPad Prism software (GraphPad Software, La Jolla, Calif) was used for data analysis. Differences between groups were determined by using the Student *t* test and 2-way ANOVA. Results are presented as means ± SEMs. Differences, indicated as 2-tailed *P* values, were considered significant at a *P* value of less than .05.

## SNP identification in human subjects

Human samples from England, Scotland, and Ireland were obtained with consent from patients and Institutional Ethics Committee approval complying with the principles of the Declaration of Helsinki. Human *MATT* was sequenced by using primers listed in Table E1, *C*, in this article's Online Repository at www.jacionline.org. Conditions for all exons except exon 2.2 were as follows: 1 cycle of 94°C for 5 minutes; 35 cycles of 94°C for 30 seconds, 54°C for 30 seconds, and 72°C for 1 minute; and a final extension at 72°C for 5 minutes. An annealing temperature of 58°C was used for exon 2.2. SNP rs6684514 (see Table E2 in this article's Online Repository at www.jacionline.org) was genotyped across the case-control study populations by using the TaqMan allelic discrimination genotyping assay (C_25986870_10, Life Technologies Corporation).

## Case-control studies

Five independently recruited AD case collections were compared with ethnically matched population control subjects (see the Methods section in this article's Online Repository). Case definitions are described in the Methods section in this article's Online Repository, and demographic data relating to these AD case collections and control subjects are summarized in Table E3 in this article's Online Repository at www.jacionline.org.

## Genotype imputation

For *in silico* replication, we used the 2 German AD case-control selections (see Table E3), which had been genotyped on Illumina 300k and Affymetrix

6.0p (Affymetrix, Santa Clara, Calif), respectively, and imputed based on the 1000 Genomes database (phase I integrated variant set release, March 2012). Previous imputation data sets were curated, and samples with a greater than 5% missing rate, excess of heterozygosity (±5 SDs away from the sample mean), and unexpected relatedness (PI_HAT>0.1875, halfway between expected IBD for third- and second-degree relatives) were excluded. Further outliers from the multidimensional scaling of the pairwise IBS matrix were removed. Imputation was carried out with SHAPEIT and IMPUTE2. The SNP rs6684514 investigated *in silico* had an imputation quality of 0.996 (proper info) and a CR of 0.997.

### Statistical genetics

Case-control comparisons for each case collection and population control subjects were performed by using logistic regression analysis in Stata 12.0 (StataCorp, College Station, Tex). Meta-analysis (under fixed- and random-effects models) of the resulting log ORs and their SEs was performed with the "metan" function in Stata 12.0.

## RESULTS

### Identification of *Matt* as the gene mutation in *matted* mice

To separate the 2 mutations in the DM *maft* mice, we used the backcross-intercross protocol outlined in Fig E1. Mice with the $Flg^{ft}$ mutant allele (*Flg* c.5303delA) were genotyped, as previously described.[12] Progeny from putative heterozygous intercrosses were checked for the recessive matted hair phenotype to identify *ma/ma* homozygotes. By this means, we separated both alleles. Because SNP analysis showed that the initial DM *maft* mouse was maintained on a mixed-strain background, we generated single-mutant N12 mice congenic on the C57BL/6J background. By using this backcross-intercross strategy to generate congenic *ma/ma* homozygous mice (see Fig E1), the recessive disease–causing locus should be ultimately visible as a small area of homozygosity flanked by 2 unequal heterozygous regions (see Fig E2, *A*, in this article's Online Repository at www.jacionline.org). To map this, we sequenced several short PCR fragments across the region from C57BL/6J congenic *ma/ma* homozygotes (see Table E1, *A*) to identify SNPs. By this means, a 1.05-Mb critical interval was defined (Fig 1, *A*). In parallel, we used epidermal transcriptomic sequencing to compare WT C57BL/6J and congenic *ma/ma* homozygotes. A nonsense mutation, p.Y280*, in the *Tmem79* gene, located within the interval (Fig 1, *A*), was identified in the RNAseq data (not shown). The mutation in *Tmem79* was confirmed by using conventional sequencing of genomic PCR products (Fig 1, *B*) and restriction digests (see Fig E2, *B*). This mutant transcript was expressed at approximately 30% of the WT levels in *ma/ma* epidermis (Fig 1, *C*). By using a commercially available antibody against the C-terminus of the encoded protein, a single band of approximately 49 kDa was seen in epidermal extracts from WT and congenic $Flg^{ft/ft}$ mice, with the protein absent in congenic *ma/ma* and DM mice (Fig 1, *D*), indicating the presence of a truncated protein or an absence of protein in these mice. Interestingly, epidermal extracts from $Flg^{ft/ft}$ mice had increased expression of protein relative to WT extracts. *Tmem79* was renamed *Matt*, encoding the protein mattrin; similarly, the human ortholog was renamed *MATT* (approved by the Genome Nomenclature Committee). By convention, after identification of the gene, the *ma* mutant allele was redesignated as $Matt^{ma}$.

### Mattrin is expressed in mouse and human skin

*Matt* is a compact 5-exon gene on mouse 3qF1 spanning 5781 bp of genomic DNA, encoding a 391-amino-acid protein with a calculated molecular weight of 43.5 kDa. The human ortholog, *MATT*, which is syntenic on 1q23.1, has an identical organization. Mattrin is a transmembrane domain protein with no previously ascribed function (Fig 1, *E*). Bioinformatics analysis predicts a long internal N-terminal domain, 5 transmembrane domains, and a short C-terminus (Fig 1, *E*, and see Fig E3 in this article's Online Repository at www.jacionline.org). The mutation p.Y280* occurs before the third transmembrane domain (Fig 1, *E*). Immunohistochemistry revealed that mattrin was present in epidermal granular layer keratinocytes in WT mice (Fig 2, *A*) but was absent from $Matt^{ma/ma}$ animals (Fig 2, *B*). Similarly, the protein was observed in granular layer cells in human epidermis (see Fig E4, *A* and *B*, in this article's Online Repository at www.jacionline.org) and was also found in the hair follicles of mice (Fig 2, *A*) and human subjects (see Fig E4, *C* and *D*). In addition, the mRNA is widely expressed in human tissues and, consistent with immunohistochemistry, is strongly expressed in skin (see Fig E4, *E*).

Expression quantitative trait locus analysis[23] showed that *Matt* is in a network of proteins expressed late in epidermal differentiation, with an expression quantitative trait locus profile closely matching that of *Rhbg* (transporter protein) and *Rab25* (membrane trafficking, data not shown). Mattrin shows distant sequence homology to the Membrane-Associated Proteins in Eicosanoid and Glutathione metabolism (MAPEG) protein family (see Fig E3),[24] members of which have roles in lipid catabolism.[25] Immunohistochemistry with the lipophilic dye Nile Red revealed highly organized stacks of cornified cell envelopes in the stratum corneum of WT animals (Fig 2, *C*) that were highly disorganized in $Matt^{ma/ma}$ mice, with discontinuous, uneven, and highly disorganized cornified cell envelopes (Fig 2, *D*).

### $Matt^{ma/ma}$ mice have spontaneous dermatitis

Gross examination of adult $Flg^{ft/ft}$ and $Matt^{ma/ma}$ mice demonstrates that the 2 strains differ considerably from each other (Fig 3, *A*). Macroscopic clinical scoring[22] demonstrated $Matt^{ma/ma}$ and DM mice spontaneously having progressive dermatitis-like skin inflammation, which did not occur in adult $Flg^{ft/ft}$ animals (see Fig E5, *A*, in this article's Online Repository at www.jacionline.org). Although all $Matt^{ma/ma}$ and DM mice had marked skin inflammation, there was a broad spectrum of pathology, with some animals exhibiting profound lesions and excoriation and occasional blepharitis and eyelid dermatitis (see Fig E6, *B*, in this article's Online Repository at www.jacionline.org). Interestingly, both neonatal $Flg^{ft/ft}$ and DM mice have significant ichthyosis ($P < .001$) compared with WT mice, but $Matt^{ma/ma}$ mice did not, indicating the postnatal importance of filaggrin (see Fig E5, *A*, and E6, *A*).

With respect to the matted hair phenotype,[11,26] both $Matt^{ma/ma}$ and DM animals had the keratinization defect, with hairs forming clumps[11,26] and hair breakage and occasional alopecia evident at 32 weeks (Fig 3, *A*, and see Fig E6, *B*). Scanning electron microscopy confirmed $Matt^{ma/ma}$ and DM mice had fragile hairs prone to longitudinal splitting and breakage with defective cuticle morphology (see Fig E7, *A*, in this article's Online Repository at www.jacionline.org). Furthermore, $Matt^{ma/ma}$ mice had distorted hair follicle morphogenesis (see Fig E7, *B*).

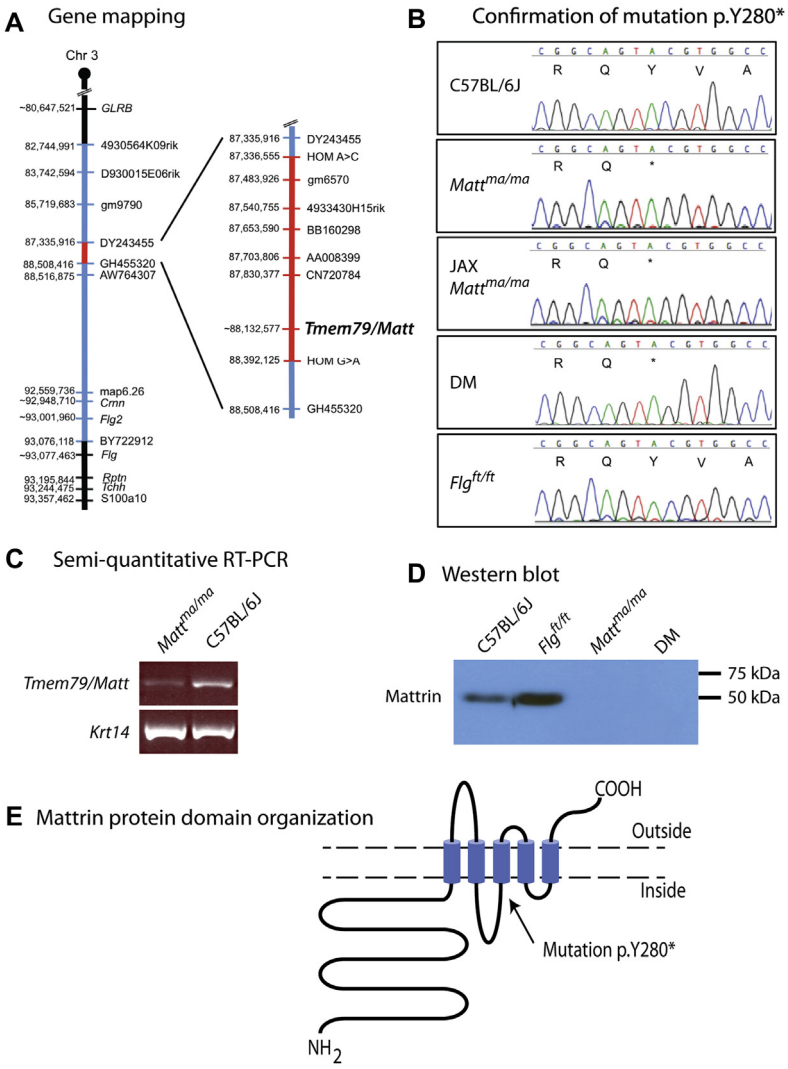



**FIG 1.** Identification of *Tmem79/Matt* as the *matted* gene. **A,** A region of homozygous *ma/ma* on chromosome 3 indicated by regions in *red* and heterozygous regions in *blue*, with the homozygous C57BL6/J genome shown in *black*. **B,** The *ma* mutation in *Tmem79/Matt* identified by means of epidermal transcriptome sequencing was confirmed by using Sanger sequencing, causing a C to G substitution at position 88,136,485 on chromosome 3. This predicts the protein change p.Y280X in the *Tmem79* (mattrin) protein. The mutant allele was renamed *Matt*$^{ma}$. **C,** Semiquantitative RT-PCR of epidermal mRNA showed a reduction in expression of *Matt* in *Matt*$^{ma/ma}$ animals. **D,** Mattrin detected in immunoblotting of skin from WT and *Flg*$^{ft/ft}$ mice but no protein expression in *Matt*$^{ma/ma}$ and DM mice. **E,** Bioinformatics shows mattrin consists of 5 transmembrane helices with a long intracellular N-terminus and short extracellular C-terminus.

Skin histology (Fig 3, *B*) showed that *Matt*$^{ma/ma}$ and DM mice have marked acanthosis ($P < .001$) and prominent orthokeratosis ($P < .001$) with dermal inflammatory infiltrates ($P < .001$; see Fig E5, *B*). In contrast, *Flg*$^{ft/ft}$ mice had occasional foci of acanthosis, mild diffuse orthokeratosis, and increased dermal cell infiltration relative to WT mice ($P < .001$). However, the phenotype in *Flg*$^{ft/ft}$ mice was subclinical, with no overt dermatitis (Fig 3, *A*, and see Fig E5, *A*).

## *Matt*$^{ma/ma}$ mutant mice are atopic and have a defective skin barrier

All mutant strains spontaneously had increased IgE levels relative to WT mice, which was markedly more pronounced in *Matt*$^{ma/ma}$ and DM animals relative to *Flg*$^{ft/ft}$ mice ($P < .001$; see Fig E5, *D*). Because the severity of skin inflammation in AD parallels barrier permeability,[27,28] TEWL was analyzed to quantify skin barrier dysregulation in mouse strains. TEWL was significantly increased in adult *Matt*$^{ma/ma}$ and DM mice ($P < .001$; see Fig E5, *C*). In contrast, TEWL was unaltered in *Flg*$^{ft/ft}$ mice relative to that seen in WT animals (see Fig E5, *C*).

HDM allergen was applied to the intact skin of *Matt*$^{ma/ma}$ and WT mice to address whether the altered skin barrier in *Matt*$^{ma/ma}$ mice influences allergen sensitization. Percutaneous sensitization with allergen led to enhanced skin inflammation in *Matt*$^{ma/ma}$ mice (Fig 4, *A-C*). Consistent with the defective barrier, TEWL was significantly upregulated ($P < .001$) in HDM-treated *Matt*$^{ma/ma}$ mice after allergen application (Fig 4, *D*). Crucially,

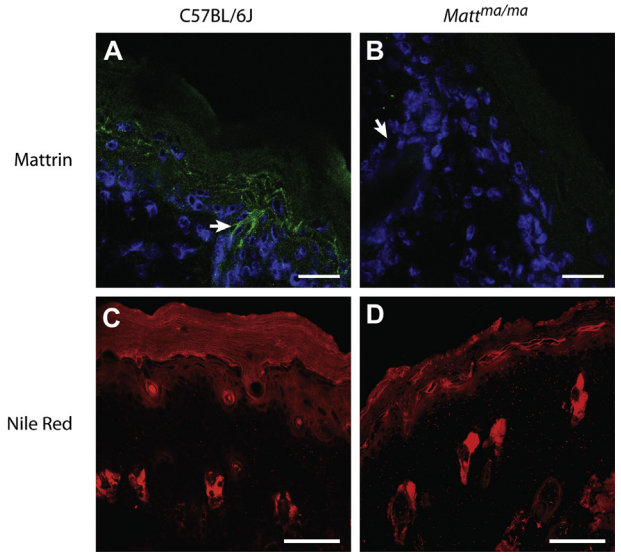



**FIG 2.** Mattrin expression and disorganized lipid morphology in the epidermis of *Matt*$^{ma}$ mice. **A** and **B,** Immunohistochemistry detection of mattrin expression confined to the epidermal granular layer of WT (Fig 2, *A*) but not *Matt*$^{ma/ma}$ (Fig 2, *B*) animals, including hair follicles *(arrow). Scale bar* = 20 μm. **C** and **D,** Nile Red staining of lipids in the epidermis of WT (Fig 2, *C*) and *Matt*$^{ma/ma}$ (Fig 2, *D*) mice revealed uneven and highly disorganized cornified cell envelopes in the stratum corneum of *Matt*$^{ma/ma}$ mice. *Scale bar* = 50 μm.

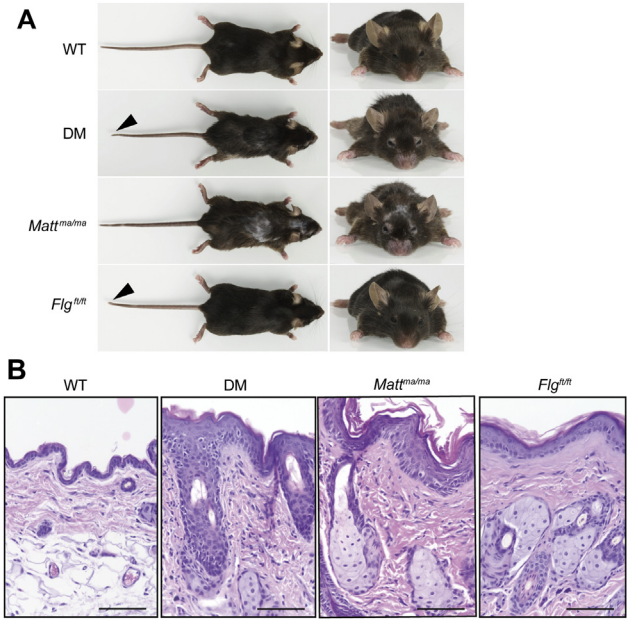



**FIG 3.** *Matt*$^{ma}$ mice are phenotypically distinct from *Flg*$^{ft/ft}$ mice, with histopathology demonstrating enhanced cutaneous inflammation at 32 weeks. **A,** Gross phenotype of WT, DM, *Matt*$^{ma/ma}$, and *Flg*$^{ft/ft}$ mice. *Flg*$^{ft/ft}$ mice are indistinguishable from WT mice, apart from the stubbed tail and shortened ear pinnae, which are phenotypic features shared with the DM mouse. All mice are age-matched homozygous males. **B,** Representative skin biopsy specimens from WT, DM, *Matt*$^{ma/ma}$, and *Flg*$^{ft/ft}$ mice at 32 weeks, showing markedly increased cutaneous inflammation in DM and *Matt*$^{ma/ma}$ mice in comparison with that seen in WT and *Flg*$^{ft/ft}$ mice. *Scale bar* = 50 μm.

HDM-specific IgE responses were significantly ($P < .05$) evoked in the sera of HDM-treated *Matt*$^{ma/ma}$ mice, as well as HDM-specific $IgG_1$ and $IgG_{2a}$ responses (Fig 4, *E*). In contrast, the skin of WT (Fig 4) and *Flg*$^{ft/ft}$ (data not shown) mice was refractory to cutaneous HDM exposure. Collectively, these data demonstrate that *Matt*$^{ma/ma}$ mice have spontaneous AD and a defective skin barrier that facilitates percutaneous allergic sensitization.

## A mutation in *MATT* is associated with AD

To investigate whether *MATT* is associated with human AD, all exons of human *MATT* were sequenced in 55 Irish AD cases who were negative for *FLG* null mutations. Several noncoding SNPs and 1 known missense SNP, rs6684514, were identified (see Table E2 in this article's Online Repository at www.jacionline.org). Case-control analyses conducted on 2 independent AD case collections, English adult AD and United Kingdom pediatric AD, with separate population-matched control subjects (see Table E3) showed a significant association between rs6684514 and AD (Table I). The minor allele (A) is protective for disease in both the English adult AD (OR, approximately 0.791; $P = .0038$) and UK pediatric AD (OR, approximately 0.770; $P = .0143$) populations (Table I). Notably, when controlling for the strong and significant *FLG* null mutations, it was shown that the association of rs6684514 with AD is independent of *FLG* (Table I and see Table E4 in this article's Online Repository at www.jacionline.org). The association of rs6684514 with AD was further replicated in a German AD case-control analysis in which the OR after controlling for *FLG* mutations was 0.86 (95% CI, 0.76-0.97; $P = .0161$; Table I). However, this association was not replicated in Irish or Scottish case-control analyses (Table I). A meta-analysis using a fixed-effects model of all 4,245 AD cases from England, Scotland, Ireland, and Germany with 10,558 population-matched control subjects did not show significant heterogeneity between the study groups ($P = .078$). A small but significant effect of rs6684514 on AD risk was confirmed in the meta-analysis (OR, 0.91; 95% CI, 0.86-0.96; $P = .001$; Fig 5; see Table E4). A random-effects meta-analysis also produced similar results (OR, 0.90; 95% CI, 0.82-0.98; $P = .015$; see Fig E8 in this article's Online Repository at www.jacionline.org).

## DISCUSSION

Previously, we demonstrated that *FLG* loss-of-function mutations have a very strong relevance for the common inflammatory skin disease AD and associated atopic phenotypes[3,9,10] and identified an analogous mutation in the murine homolog *Flg* in the spontaneously occurring flaky tail DM (*maft*) mouse.[12] Subsequently, these DM mice have been widely used as a model of heritable skin barrier deficiency and spontaneous dermatitis[13-15,17] and as a model of filaggrin deficiency–associated AD pathogenesis in patients.[29] In this study we have identified the *matted* phenotype in the DM mouse as arising from a second mutation in *Matt* (*Tmem79*). The *Matt*$^{ma}$ mutation results in defective expression of the transmembrane protein mattrin, which is highly expressed in the upper granular layer of epidermal keratinocytes, with a predicted role in lipid homeostasis. These studies have elucidated the relative contributions of the *Flg* and *Matt* mutations in the DM mouse. The *Matt*$^{ma}$ mutation led to the striking development of spontaneous AD-like skin pathology and atopy in adult mice. These data indicate that the DM mouse is not a true model of filaggrin deficiency–associated AD-like skin

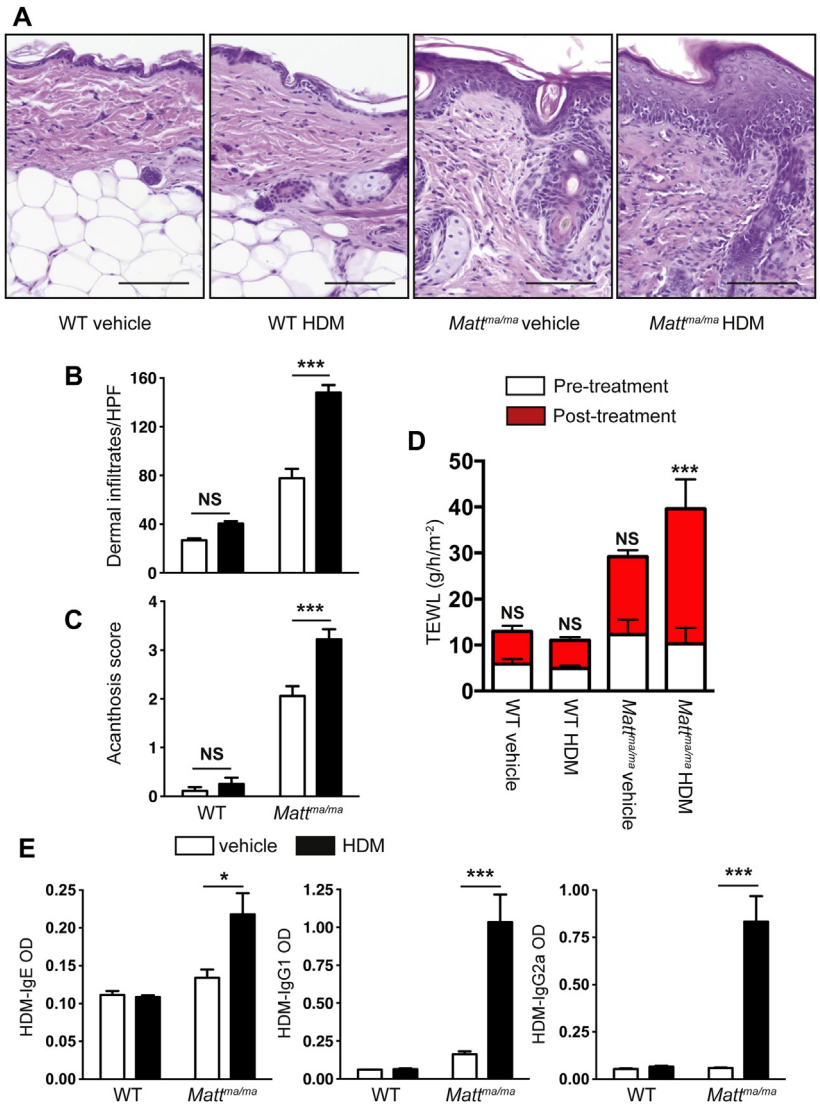



**FIG 4.** *Matt*$^{ma}$ mice have exacerbated AD-like inflammation and HDM-specific responses after allergen challenge to intact skin. **A,** Representative skin biopsy specimens from age-matched WT and *Matt*$^{ma/ma}$ mice treated with HDM or vehicle, with increased inflammation in HDM-treated *Matt*$^{ma/ma}$ mice relative to that seen in HDM-treated WT and vehicle-treated *Matt*$^{ma/ma}$ mice. *Scale bar* = 50 μm. **B,** HDM treatment induces increased dermal cell infiltration in *Matt*$^{ma/ma}$ mice relative to that seen in HDM-treated WT and vehicle-treated *Matt*$^{ma/ma}$ mice. **C,** Acanthosis is increased in HDM-treated *Matt*$^{ma/ma}$ mice relative to that seen in HDM-treated WT and vehicle-treated *Matt*$^{ma/ma}$ mice. **D,** Epicutaneous HDM treatment results in significantly increased TEWL in *Matt*$^{ma/ma}$ mice relative to that seen in HDM-treated WT mice. **E,** HDM-treated *Matt*$^{ma/ma}$ mice have increased HDM-specific serum IgE, $IgG_1$, and $IgG_2$ levels relative to those seen in HDM-treated WT mice. Cell numbers and acanthosis were scored on 15 to 20 hpfs (×1000 magnification) on hematoxylin and eosin–stained sections from WT, DM, *Matt*$^{ma/ma}$, and *Flg*$^{ft/ft}$ mice. Data represent the mean and *error bars* represent ± SEMs from 6 to 8 mice and are representative of 2 separate experiments. The Student *t* test or 2-way ANOVA was used to determine statistical differences between groups. *NS*, Not significant, **P* > .05. ****P* < .001.

inflammation. It is notable that the *Flg*$^{ft}$ mutation has a neonatal influence, whereas the *Matt*$^{ma}$ mutation has a progressive influence with age. This indicates the polygenic nature of the DM mouse as an AD model and indicates the differential influence of both the *Flg*$^{ft}$ and *Matt*$^{ma}$ mutations. Recently, *Flg*$^{-/-}$ mice were shown to have no overt dermatitis with age,[30] with the authors unable to sensitize *Flg*$^{-/-}$ mice by means of application of OVA, an allergen commonly used in models of skin inflammation,[31] to the intact skin barrier, results similar to our data. Notably, *Flg*$^{ft/ft}$ mice were not sensitized to HDM (data not shown), indicating an impermissibility of the *Flg*$^{ft/ft}$ skin barrier to protein antigen ingress. These data indicate that the DM mouse is not a true model of filaggrin deficiency–associated AD-like skin inflammation.

In this study we adopted a translational approach and also addressed whether mattrin was implicated in human AD. A missense SNP in *MATT* was shown to have a small but significant effect on the risk for human AD. Because mattrin shows

**TABLE I.** Results of case-control analyses to investigate the association of rs6684514 and AD in 5 populations

| Case-control comparison | Cases (n) | Control subjects (n) | *P* value for rs6684514 | OR | 95% CI | *P* value for rs6684514 association after controlling for *FLG* null mutations |
|---|---|---|---|---|---|---|
| English adult severe AD vs English population control subjects from the 1958 Birth Cohort | 505 | 1919 | .0038 | 0.791 | 0.674-0.929 | .0008 |
| UK mild-moderate pediatric AD vs English pediatric control subjects without AD | 338 | 538 | .0153 | 0.770 | 0.622-0.953 | .0328 |
| Irish pediatric AD vs Irish adult population control subjects | 724 | 1905 | .1300 | 1.025 | 0.896-1.172 | .7905 |
| German AD cases vs German population control subjects | 1543 | 2005 | .0794 | 0.912 | 0.822-1.011 | .0161 |
| Scottish asthma cases with AD vs Scottish population control subjects | 1135 | 4189 | .1365 | 0.922 | 0.831-1.023 | NA |

*English adult severe AD* is defined as early-onset, persistent, and severe disease. *UK mild-moderate pediatric AD* includes cases collected from an English population birth cohort (n = 177) and a Scottish General Practice collection (n = 161); the *English pediatric control subjects* were ascertained not to have AD at the age of 7 to 9 years. *Irish pediatric AD* cases were collected in secondary and tertiary care clinics in Ireland; *Irish population control subjects* are healthy adult blood donors. *Scottish asthma cases with AD* are subjects with physician-diagnosed asthma and parent-reported AD. The German case and control rs6684514 genotypes were ascertained by imputation from genome-wide SNP analysis data; all other SNP genotypes and *FLG* null mutations were analyzed by using TaqMan allelic discrimination assays. In the English, United Kingdom, and Irish collections, 4 prevalent *FLG* null mutations were analyzed (R501X, 2282del4, R2447X, and S3247X), whereas in the German population 2 *FLG* null mutations were analyzed (R501X and 2282del4). Statistical analysis was performed through logistic regression in Stata 10.0 software (StataCorp).
*NA*, *FLG* genotype data not available.

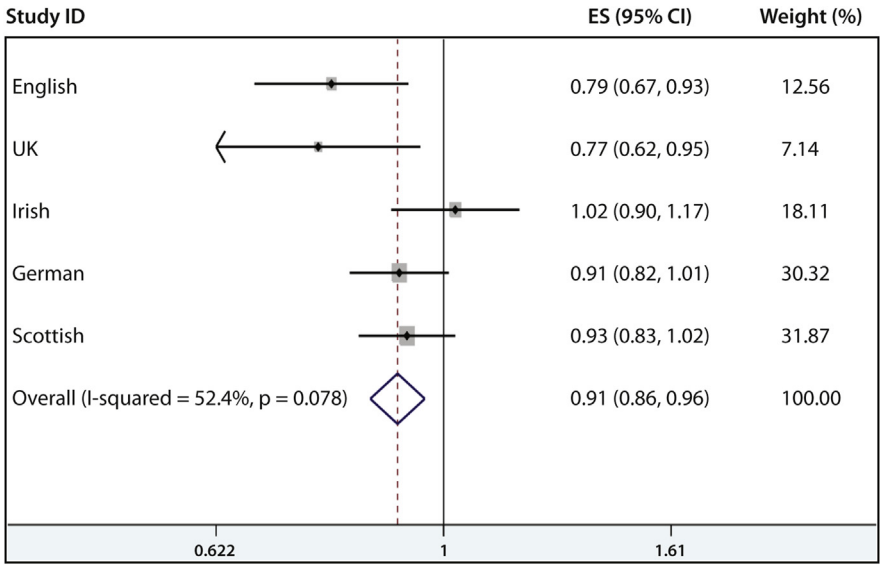



**FIG 5.** Forrest plot showing results of a fixed-effects meta-analysis of 5 case-control studies to investigate the association of rs6684514 and AD. Study populations used were as follows: *English*, English adult severe AD versus English population control subjects from the 1958 Birth Cohort; *UK*, UK mild-moderate pediatric AD versus English pediatric control subjects without AD; *Irish*, Irish pediatric AD versus Irish adult population controls; *German*, German AD cases versus German population control subjects; *Scottish*, Scottish asthma cases with AD versus Scottish population control subjects. *ES*, Estimated odds ratio. Meta-analysis was carried out with the "metan" function in Stata software (StataCorp).

distant sequence homology to MAPEG family members,[24] which are known to catalyze glutathione-dependent transformations of lipophilic substrates at the lipid bilayer, sequence homology suggests a possible role for mattrin in the biology of lipids or lipid-like molecules.[25] Epidermal transcriptomic analysis of WT versus *Matt*$^{ma/ma}$ mice has revealed changes in several transcripts involved in fatty acid and lipid biology (data not shown). However, 1-dimensional and 2-dimensional thin-layer chromatography of epidermal lipid extracts did not reveal any differences in the migration or abundance of ceramides, phospholipids, or polar or nonpolar lipids between WT and *Matt*$^{ma/ma}$ mice (data not shown). Future work will require analysis of the glutathione binding of mattrin, elucidation of lipophilic substrates of mattrin, and an investigation of the downstream signaling pathways and local immunologic milieu in *Matt*$^{ma}$ mice.

In summary, using a translational mouse-patient strategy, we have identified a new gene mutation that leads to dermatitis in mice and have further demonstrated that a variant in the human gene is associated with AD.

We thank the patients affected by eczema and their families for their interest and participation. We also thank Masa Amagai and his research group from Keio University, Tokyo, Japan, for their open and enthusiastic collaboration, discussion, and sharing of data.

**Clinical implications: The role of *MATT* mutations in susceptibility to AD and related allergic conditions should be further investigated.**

## REFERENCES

1. Deckers IA, McLean S, Linssen S, Mommers M, van Schayck CP, Sheikh A. Investigating international time trends in the incidence and prevalence of atopic eczema 1990-2010: a systematic review of epidemiological studies. PLoS One 2012;7:e39803.
2. Spergel JM, Paller AS. Atopic dermatitis and the atopic march. J Allergy Clin Immunol 2003;112(Suppl):S118-27.
3. Irvine AD, McLean WH, Leung DY. Filaggrin mutations associated with skin and allergic diseases. N Engl J Med 2011;365:1315-27.
4. Esparza-Gordillo J, Weidinger S, Folster-Holst R, Bauerfeind A, Ruschendorf F, Patone G, et al. A common variant on chromosome 11q13 is associated with atopic dermatitis. Nat Genet 2009;41:596-601.
5. Hirota T, Takahashi A, Kubo M, Tsunoda T, Tomita K, Sakashita M, et al. Genome-wide association study identifies eight new susceptibility loci for atopic dermatitis in the Japanese population. Nat Genet 2012;44:1222-6.
6. Paternoster L, Standl M, Chen CM, Ramasamy A, Bonnelykke K, Duijts L, et al. Meta-analysis of genome-wide association studies identifies three new risk loci for atopic dermatitis. Nat Genet 2012;44:187-92.
7. Sun LD, Xiao FL, Li Y, Zhou WM, Tang HY, Tang XF, et al. Genome-wide association study identifies two new susceptibility loci for atopic dermatitis in the Chinese Han population. Nat Genet 2011;43:690-4.
8. Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, Zhao Y, et al. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. Nat Genet 2006;38:337-42.
9. Palmer CN, Irvine AD, Terron-Kwiatkowski A, Zhao Y, Liao H, Lee SP, et al. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. Nat Genet 2006;38:441-6.
10. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM, et al. Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. Nat Genet 2007;39:650-4.
11. Presland RB, Boggess D, Lewis SP, Hull C, Fleckman P, Sundberg JP. Loss of normal profilaggrin and filaggrin in flaky tail (ft/ft) mice: an animal model for the filaggrin-deficient skin disease ichthyosis vulgaris. J Invest Dermatol 2000;115:1072-81.
12. Fallon PG, Sasaki T, Sandilands A, Campbell LE, Saunders SP, Mangan NE, et al. A homozygous frameshift mutation in the mouse Flg gene facilitates enhanced percutaneous allergen priming. Nat Genet 2009;41:602-8.
13. Moniaga CS, Egawa G, Kawasaki H, Hara-Chikuma M, Honda T, Tanizaki H, et al. Flaky tail mouse denotes human atopic dermatitis in the steady state and by topical application with *Dermatophagoides pteronyssinus* extract. Am J Pathol 2010;176:2385-93.
14. Oyoshi MK, Murphy GF, Geha RS. Filaggrin-deficient mice exhibit TH17-dominated skin inflammation and permissiveness to epicutaneous sensitization with protein antigen. J Allergy Clin Immunol 2009;124:485-93.e1.
15. Scharschmidt TC, Man MQ, Hatano Y, Crumrine D, Gunathilake R, Sundberg JP, et al. Filaggrin deficiency confers a paracellular barrier abnormality that reduces inflammatory thresholds to irritants and haptens. J Allergy Clin Immunol 2009;124:496-506, e1-6.
16. Sundberg JP. The flaky tail (ft) mutation. Handbook of mouse mutations with skin and hair abnormalities animal models and biochemical tools. CRC press:1984;2:269-73.
17. Moniaga CS, Jeong SK, Egawa G, Nakajima S, Hara-Chikuma M, Jeon JE, et al. Protease activity enhances production of thymic stromal lymphopoietin and basophil accumulation in flaky tail mice. Am J Pathol 2013;182:841-51.
18. de Guzman Strong C, Conlan S, Deming CB, Cheng J, Sears KE, Segre JA. A milieu of regulatory elements in the epidermal differentiation complex syntenic block: implications for atopic dermatitis and psoriasis. Hum Mol Genet 2010;19:1453-60.
19. Kizawa K, Ito M. Characterization of epithelial cells in the hair follicle with S100 proteins. Methods Mol Biol 2005;289:209-22.
20. Kizawa K, Takahara H, Unno M, Heizmann CW. S100 and S100 fused-type protein families in epidermal maturation with special focus on S100A3 in mammalian hair cuticles. Biochimie 2011;93:2038-47.
21. O'Keefe EJ, Hamilton EH, Lee SC, Steinert P. Trichohyalin: a structural protein of hair, tongue, nail, and epidermis. J Invest Dermatol 1993;101(Suppl):65S-71S.
22. Rice RH, Wong VJ, Pinkerton KE, Sundberg JP. Cross-linked features of mouse pelage hair resistant to detergent extraction. Anat Rec 1999;254:231-7.
23. Quigley DA, To MD, Perez-Losada J, Pelorosso FG, Mao JH, Nagase H, et al. Genetic architecture of mouse skin inflammation and tumour susceptibility. Nature 2009;458:505-8.
24. Jakobsson PJ, Morgenstern R, Mancini J, Ford-Hutchinson A, Persson B. Membrane-associated proteins in eicosanoid and glutathione metabolism (MAPEG). A widespread protein superfamily. Am J Respir Crit Care Med 2000;161(Suppl):S20-4.
25. Martinez Molina D, Eshaghi S, Nordlund P. Catalysis within the lipid bilayer-structure and mechanism of the MAPEG family of integral membrane proteins. Curr Opin Struct Biol 2008;18:442-9.
26. Chamlin SL, Kao J, Frieden IJ, Sheu MY, Fowler AJ, Fluhr JW, et al. Ceramide-dominant barrier repair lipids alleviate childhood atopic dermatitis: changes in barrier function provide a sensitive indicator of disease activity. J Am Acad Dermatol 2002;47:198-208.
27. Gupta J, Grube E, Ericksen MB, Stevenson MD, Lucky AW, Sheth AP, et al. Intrinsically defective skin barrier function in children with atopic dermatitis correlates with disease severity. J Allergy Clin Immunol 2008;121:725-30, e2.
28. Sugarman JL, Fluhr JW, Fowler AJ, Bruckner T, Diepgen TL, Williams ML. The objective severity assessment of atopic dermatitis score: an objective measure using permeability barrier function and stratum corneum hydration with computer-assisted estimates for extent of disease. Arch Dermatol 2003;139:1417-22.
29. Moniaga CS, Kabashima K. Filaggrin in atopic dermatitis: flaky tail mice as a novel model for developing drug targets in atopic dermatitis. Inflamm Allergy Drug Targets 2011;10:477-85.
30. Kawasaki H, Nagao K, Kubo A, Hata T, Shimizu A, Mizuno H, et al. Altered stratum corneum barrier and enhanced percutaneous immune responses in filaggrin-null mice. J Allergy Clin Immunol 2012;129:1538-46, e6.
31. Oyoshi MK, He R, Li Y, Mondal S, Yoon J, Afshar R, et al. Leukotriene B4-driven neutrophil recruitment to the skin is essential for allergic skin inflammation. Immunity 2012;37:747-58.